THREE-DIMENSIONAL DISTRIBUTED EFFECTS ON THE DC AND AC PERFORMANCE OF HIGH SPEED BIPOLAR TRANSISTORS

BY

MING-YEH CHUANG

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1997

ACKNOWLEDGEMENTS

I wish to express my deepest gratitude to my advisor, Dr. Mark E. Law, whose constant support and patient guidance provided a clear path for my research. I thank Dr. Kenneth K. O. with whom I had many stimulating discussions. I would also like to acknowledge Dr. Jerry G. Fossum, Dr. Sheng S. Li, Dr. Robert M. Fox, and Dr. Loc Vu-Quoc for their time and advice and for serving on my supervisory committee. Thanks also go to Dr. Rex Lowther at Harris Semiconductor for providing the BJT device information, which initiated the study of the distributed effects.

I have been quite fortunate to have worked with my colleagues. Susan Earles spent a long time in correcting my writing and speech. Hernan Rueda provided assistance in 3D mesh. Numerous ideas have grown out of my discussions with Jon Brodsky, Dukhyun Chang, Samir Chaudhry, Chih-Ming Hung, and Meng-Hsueh Chiang.

I am most pleased to acknowledge the support and encouragement of my parents and my family to whom I dedicate this work.

TABLE OF CONTENTS

page

ACKNOWLEDGEMENTS ii

ABSTRACT vi

CHAPTERS

1 INTRODUCTION 1

2 IMPLEMENTATION OF NEW PHYSICAL MODELS AND FUNCTIONS IN FLOODS 6

2.1 Introduction 6
2.2 Single Node Self-Heating Model 7
2.2.1 Motivation 7
2.2.2 Model Development 8
2.2.3 Results and Discussion 12
2.3 Current Boundary Condition 18
2.4 Interface Transport Model of Polysilicon Emitter Contacts 21
2.4.1 Motivation 21
2.4.2 Model Implementation 23
2.4.3 Results and Discussion 26
2.5 Device/Circuit Tightly Coupled Simulator 31
2.5.1 Motivation 31
2.5.2 Algorithms 31
2.5.3 D.C. Simulations 37
2.5.4 A.C. Simulations 47
2.6 Summary 48

3 THREE DIMENSIONAL BASE DISTRIBUTED EFFECTS OF STRIPE BJTS -- BASE RESISTANCE AT DC 50

3.1 Introduction 50
3.2 Asymmetric Distributed Model in the Active Base Region 51
3.2.1 Modeled Device Structure 51
3.2.2 Intrinsic Model Development 52
3.2.3 Model with Extrinsic Base Resistance 55
3.2.4 Verification and Discussion 57
3.3 Quasi Three Dimensional DC Distributed Mode 62

3.3.1 Model Development 62
3.3.2 Results and Discussion 64
3.3.3 Model Verification/Application 67
3.4 Summary 73

4 THREE DIMENSIONAL BASE DISTRIBUTED EFFECTS OF STRIPE BJTS -- AC EFFECTS ON INPUT CHARACTERISTICS 78

4.1 Introduction 78
4.2 A.C. Distributed Model in the Active Base Region 79
4.2.1 Model Development 79
4.2.2 Verification and Discussion 83
4.3 Three Dimensional A.C. Distributed Model 89
4.3.1 Model Development 89
4.3.2 Results and Discussion 93
4.3.3 Application and Verification 100
4.4 Summary 105

5 VERIFICATION OF DC DISTRIBUTED MODEL AND GEOMETRY DEPENDENT DEVCIE PERFORMANCE DUE TO DISTRIBUTED EFFECTS 107

5.1 Introduction 107
5.2 3D Device Simulations and Simulation Structures 109
5.3 D.C. Characteristics 114
5.4 Summary 124

6 VERIFICATION OF AC DISTRIBUTED MODEL AND GEOMETRY DEPENDENT AC DEVICE PERFORMANCE DUE TO DISTRIBUTED EFFECTS 127

6.1 Introduction 127
6.2 Verification of AC Model and Quasi-3D Simulations 128
6.3 Structure and Geometry Dependent AC Performance 134
6.4 The Importance of Distributed Effects on Circuit Design 145
6.5 Summary 149

7 THE DISTRIBUTED EFFECTS ON THE AC AND TRANSIENT PERFORMANCE OF A 0.18μm MOSFET 152

7.1 Introduction 152
7.2 Transient Simulations 154
7.3 Distributed Effects on the Transient Performance of MOSFETs 155
7.4 Distributed Effects on the AC Performance of MOSFETs 160
7.4 Summary 161

8 SUMMARY AND FUTURE WORK164

8.1 Summary....165
8.2 Future Work....166

REFERENCES....169

BIOGRAPHICAL SKETCH....175

Abstract of Dissertation presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

THREE-DIMENSIONAL DISTRIBUTED EFFECTS
ON THE DC AND AC PERFORMANCE OF
HIGH SPEED BIPOLAR TRANSISTORS

By Ming-Yeh Chuang

August 1997

Chairman: Dr. Mark E. Law
Major Department: Electrical and Computer Engineering

The input impedance of a transistor is usually difficult to predict either by compact models or two-dimensional (2D) device simulations due to the three-dimensional (3D) distributed effect in the input port of the device. The voltage and current distribution characteristics vary at different biases as well as at different operating frequencies due to this effect resulting in the modulation of input characteristics. In this dissertation, an analytical model describing the DC and AC voltage and current distributed effects in the polysilicon and intrinsic base regions of stripe bipolar junction transistors (BJTs) with double polysilicon technology is presented. It is shown that the bias and frequency dependent distributed effect in the base polysilicon contacts cause an unequal division of base current between the two base polysilicon contacts and result in a re-distribution of base current in the base regions at different levels of current injection and operating frequencies. Therefore, the input characteristics are also modulated at different operating conditions. It is also observed that this effect has different degrees of influence on devices with

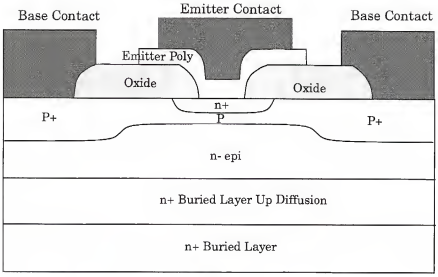
various types of layout structures, geometry sizes, and polysilicon contact resistances. Thus, resulting performance of these devices is quite different. The distributed effect on the transient and AC performance of advanced metal-oxide-semiconductor field-effect transistors (MOSFETs) is also studied based on the 3D device simulations, and the results show the distributed effect is less important for MOSFETs.

CHAPTER 1
INTRODUCTION

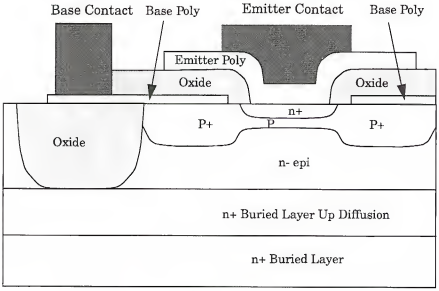
Modern high speed BJTs are built with highly scaled lateral and vertical device geometries. The vertical scaling increases the current densities and reduces the minority carrier storage whereas the lateral scaling decreases the parasitic capacitances and resistances in the devices.

In single polysilicon bipolar transistors (refer to Fig. 1.1(a)), the heavily doped polysilicon layer is used as a diffusion source to form a shallow emitter junction and serves as a low resistivity contact for the emitter. The polysilicon emitter structure has also allowed scaling of the base width while simultaneously improving the device current gain and performance [War95]. The double-polysilicon self-aligned (DPSA) bipolar device structure (refer to Fig. 1.1(b)) first appeared over ten years ago [Nak80, Tan82] and became the key technology for most high performance BJTs [Yam93, Ker92, Tor91]. The use of a heavily doped polysilicon layer as the base contact scales the extrinsic base regions and results in much smaller parasitic capacitance compared with the single polysilicon transistors [Nak95]. The double polysilicon processes with sidewall spacer technologies reduce the link between the intrinsic and extrinsic base regions and decrease the base resistance [Zde87].

The resistivity in the base region causes a debiasing effect along the distributed base current path and results in a nonuniform voltage and thus emitter current injection across the base-emitter junction. The emitter current density at the edge is higher than that at the center of the intrinsic base region



(a)



(b)

Figure 1.1 The cross section of (a) a single polysilicon and (b) a double polysilicon BJTs. Note the longer inter-connection between intrinsic and extrinsic base regions in the single polysilicon BJT.

which is the so called "current crowding" effect. Current high speed BJTs tend to operate at high current density and high frequency where the current crowding effect is prominent. The effect in the intrinsic base region on the DC and AC base resistance has been widely investigated [Yua88, Rhe95, Gho65, Ver91, Cho90]. The lateral scaling of the emitter stripe width with DPSA technologies alleviates the effect. The base polysilicon contact is also important for its role in determining the base resistance and a similar debiasing effect takes place in the polysilicon region as well. As device dimensions are scaled, the resistance through the base polysilicon contact increases since the area of the polysilicon layer must also be scaled. Thus, the debiasing effect in the polysilicon base region is even more important in the scaled DPSA transistors.

Contemporary MOSFET technologies utilize polysilicon as the gate material to reduce gate (channel) length as well as control threshold voltage. Usually the metal contacts of the polysilicon gate are attached at the end of the gate. The distributed voltage and current effects similar to that of BJTs are observed in the gate of MOSFETs at high speed operations. Shrinking the gate length greatly improves the speed and increases current density. The thickness of the gate is correspondingly scaled due to the required planarity. Both of the tendencies toward high speed operation and the scaling of feature sizes in MOSFETs enhance the distributed characteristics of the device.

Two-dimensional device simulations and compact model simulations currently are the main tools to characterize the performance of the semiconductor devices. However, the nonuniformity of voltage and current in the polysilicon contact region cannot be properly described by these methods due to the distributed voltage and current characteristics that occur in polysilicon. Therefore, the understanding and modeling of the distributed effects is important for accurate and efficient design of high speed semiconductor devices. The

goals of this work are to (i) characterize the distributed and debiasing effects, (ii) model the effects, and (iii) evaluate the impact of the effects on the DC and AC device performance.

The device simulator FLOODS (FLorida Object-Oriented Device Simulator) [Lia94a, Lia94b, Lia94c] is the main thrust for this study. To accurately describe the complicated physical mechanisms of modern semiconductor devices, several new models and functions have been added to FLOODS. The development and implementation of these functions are presented in chapter 2.

In chapter 3, a quasi-3D DC distributed model which considers the effect in both the intrinsic and polysilicon base regions is developed for BJTs and applied to the computation of base resistance. Model verification is accomplished through comparison with the measured base resistance.

In chapter 4, a quasi-3D AC distributed model is developed to analyze the current redistribution in BJTs at different operating frequencies. The model is also applied to FLOODS for quasi-3D simulations in which nonuniformity of voltage and current in the third dimension are taken into account. The model is verified by the comparison between the simulations and s-parameter measurements.

In chapter 5, a 3D device simulation scheme and device structures are presented, and the problems associated with the 3D simulations are discussed. The DC distributed effects presented in BJTs are studied with 3D simulations and compared with the results of the 3D model. The layout structure, geometry size and polysilicon sheet resistance dependence of the BJT performance due to the distributed effects are also discussed.

In chapter 6, AC distributed effects in BJTs are studies using 3D simulations. The simulation results confirm the accuracy of the AC model derived in

chapter 4. The AC performance of devices with different layout structures, geometry sizes and polysilicon sheet resistances are then characterized at different frequencies. The impacts of the distributed effects on the device and circuit design of high speed BJTs are also presented.

The implementation of a transient simulation scheme in FLOODS is presented in chapter 7. The effects of the distributed gate resistance on the transient and AC characteristics of a 0.18μm NMOS transistor with different gate sheet resistances are studied based on the results of 3D device simulations.

CHAPTER 2
IMPLEMENTATION OF NEW PHYSICAL MODELS AND FUNCTIONS IN FLOODS

2.1 Introduction

FLOODS (FLorida Object Oriented Device Simulator) [Lia94b, Lia94c] is a general purpose 3D device simulator. In addition to Poisson and two carrier continuity equations, lattice self-heating and two carrier energy balance equations can be solved in FLOODS to predict the characteristics of advanced devices with self-heating and hot carrier effects. Unlike MEDICI which decouples the energy balance equations from the carrier continuity equations, FLOODS solves the six equations simultaneously to achieve self-consistency in these equations. Another advantage of FLOODS is that it is implemented in C++ which provides benefits of modularity, great portability, and reusability of code modules.

C++ and object-oriented programming techniques are characterized by the inheritance and dynamic binding functions. Inheritance allows a new class to inherit the basic functions and data of its parents. No additional programming is required in the new class except for those functions or data that either extend or replace characteristics inherited from their parents. Dynamic binding is provided by the virtual functions: the program resolves which virtual function in the inheritance hierarchy to call based on the run-time type of the actual object. Run-time type resolution encapsulates the implementation details of an inheritance hierarchy from the programs that make use of it.

Users do not have to know how the objects are constructed and the user code is not subject to change to the hierarchy. These two features greatly simplify the implementation of new models or functions in FLOODS.

In this chapter, a single-node self-heating model is developed in FLOODS to reduce the running time of full thermodynamic simulations without losing much accuracy. The implementation of the current boundary condition is shown which provides unique and stable solution at the situations of small differential resistance or negative resistance. A carrier tunneling model in the polysilicon emitter interface is exhibited which gives a proper description of the carrier distribution and conduction in the boundary. The coupling of a circuit simulator with FLOODS is also presented to include the parasitic effects in device simulations and provide direct circuit simulations with complex device structures.

2.2 Single Node Self-Heating Model

2.2.1 Motivation

The traditional drift-diffusion (DD) transport model in device simulations solves three semiconductor equations: Poisson equation plus the electron and hole continuity equations. However, as device technologies improve, these three equations become insufficient to describe the characteristics of state-of-the-art semiconductor devices. The failure comes from the fact that a quasi-static condition is assumed while deriving the DD transport equations which is not valid for carrier transport with rapid spatial and time variation of the local electric field. The large field and field gradient inside the scaled devices cause lattice self-heating and hot carrier nonlocal transport effects. Lattice self-heating and carrier energy balance equations [Str62, Coo83, Che93, Fus95b] govern the energy transfer among electrons, holes, and the crystal lattice which

take into account most of the possible mechanisms associated with nonlocal and self-heating effects. The full thermodynamic (TD) simulations solve these three energy equations together with DD equations producing better prediction capabilities compared with the DD simulations [Lia94b, Ago94].

The specific position of a simulated device geometry is represented by a node in the mesh. Equations are solved and solutions are stored at each node. For a simulated structure of N nodes, the total equations M solved in a TD simulation are $6 \times N$ whereas $3 \times N$ equations are solved in a DD simulation. The required memory and CPU time to solve the problem are approximately proportional to $M^{1.5}$. Thus, the memory and CPU time requirements of TD simulations are 2.8 times of the DD simulations for the same number of nodes. A scaled device, in general, requires thousands of nodes to define the device structure. Although TD simulations are more accurate, the simultaneous solving of six equations is CPU and memory intensive compared with DD simulations.

In this subsection, an algorithm which reduces the solution time of the TD simulations with little sacrifice in the accuracy is presented.

2.2.2 Model Development

The energy balance equations for electrons and holes are given by

$$\frac{\partial}{\partial t}(n \cdot U_n) = -\nabla \vec{S}_n - R_{wn} \tag{2-1}$$

and

$$\frac{\partial}{\partial t}(p \cdot U_p) = -\nabla \vec{S}_p - R_{wp} \tag{2-2}$$

where n and p are the electron and hole densities, respectively, $U_{n,p}$ is the average local carrier energy, $\vec{S}_{n,p}$ is the carrier energy flux, and R_{wn} and R_{wp}

are the energy transfer rates from electron and hole, respectively, to other subsystems due to the scattering or carrier recombination/generation. The lattice self-heating equation governing joule-heating and heat conduction in the semiconductor material is

$$\frac{\partial}{\partial t}(U_L) = -\vec{\nabla} \bullet \vec{S}_L - R_{WL} \tag{2-3}$$

where U_L is the local lattice energy, $\vec{S}_L$ is the lattice heat flux, and R_{WL} is the local energy transfer rate from crystal lattice to the carrier subsystems. These three equations basically state the conservation of energy at a specific position of the device (node) during a period of time. If no external energy, for example, photon, is supplied, the complete thermodynamic system is energy conservative and we have $R_{WL} + R_{wn} + R_{wp} = 0$. The energy conservation relation and the energy flow between the three subsystems are plotted schematically in figure 2.1. The discretization scheme of eq. (2-3) with box (finite difference) method is shown in figure 2.2 and can be expressed by

$$\frac{\partial}{\partial t}(U_L) \cdot A_i = -\sum_b \vec{S}_{Lb} \cdot l_b - \sum_j \vec{S}_L \cdot l_j - R_{WL} \cdot A_i \tag{2-4}$$

where A_i is the box area of a node i, l_j is the length of the box edge between node i and bulk node j, and l_b is the length of box edge between node i and boundary b. The boundary condition of lattice self-heating is given by

$$\sum_b \vec{S}_{Lb} \cdot l_b = G_{th} \cdot (T_{L,i} - T_{L(ambient)}) \tag{2-5}$$

where G_{th} is the thermal conductance of the interconnection boundary and $T_{L,i}$ is the lattice temperature at node i.

$\vec{S}_L$ in the bulk semiconductor is primarily the heat flow due to the lattice temperature gradient and can be expressed as

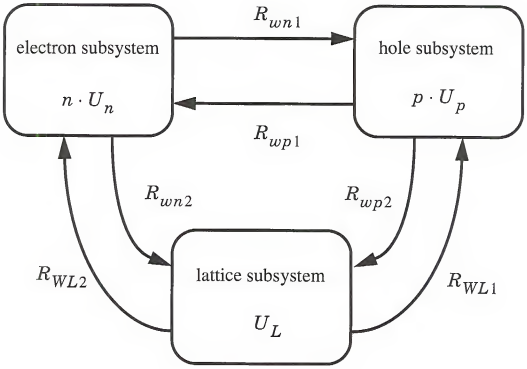


Figure 2.1. Schematic diagram of the thermodynamic system. The energy system are divided into three subsystems: electrons, holes and lattice. Energy in each subsystem can transfer to the other two subsystems by way of scattering, recombination or generation.

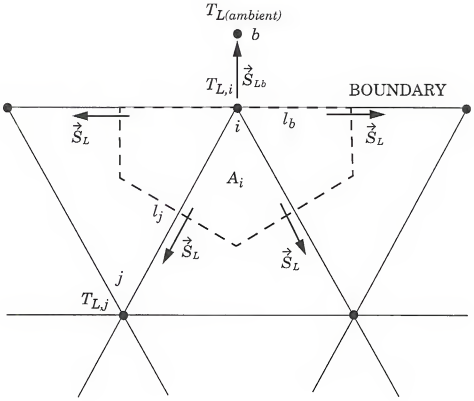


Figure 2.2. The discretization scheme with a box method in a two dimensional mesh. The shaded area indicates the area A_i associated with node i. In single node self-heating model, $T_{L,i} = T_{L,j}$.

$$\vec{S}_L = -K_L \bullet \vec{\nabla} T_L \tag{2-6}$$

where K_L is the lattice's thermal conductivity. The active region of a scaled BJT or MOS transistor is usually small and the local self-heating of the silicon lattice can be distributed over the active region. If we assume the lattice temperature is constant in the bulk region, we have $\vec{\nabla} T_L = 0$ and $\vec{S}_L = 0$ and eq. (2-4) becomes

$$\frac{\partial}{\partial t}(U_L) \cdot A_i = -\sum_b \vec{S}_{Lb} \cdot l_b - R_{WL} \cdot A_i \tag{2-7}$$

Since the lattice temperature is constant in the bulk, the solution of T_L in one node can represent T_Ls at the other N-1 nodes and only one instead of N lattice self-heating equations needs to be solved in the bulk mesh of N nodes. The size of the Jacobian matrix becomes $(5N+1)\times(5N+1)$ whereas the TD simulation is required to solve a $6N \times 6N$ matrix. Though solving one equation, FLOODS loops over the N nodes and sums the energy transfer rate R_{WL} of each node in the same equation. Thus, eq. (2-7) can be interpreted as follows: the total lattice energy generated in the device is represented by a single temperature in one node and balanced by the heat flux at the contact boundaries. In the isothermal model, the self-heating equation is not taken into account and the size of the Jacobian matrix is $5N \times 5N$ which is about the same size as the matrix in the single node self-heating (SNSH) model if N >> 1.

2.2.3 Results and Discussion

The lattice temperature distribution in an npn BJT with TD simulation of FLOODS is shown in figure 2.3. The base-emitter junction is forward biased at 1.0V which drives the device into the high current injection regime and pushs the base-collector space-charge region to the edge of the collector buried

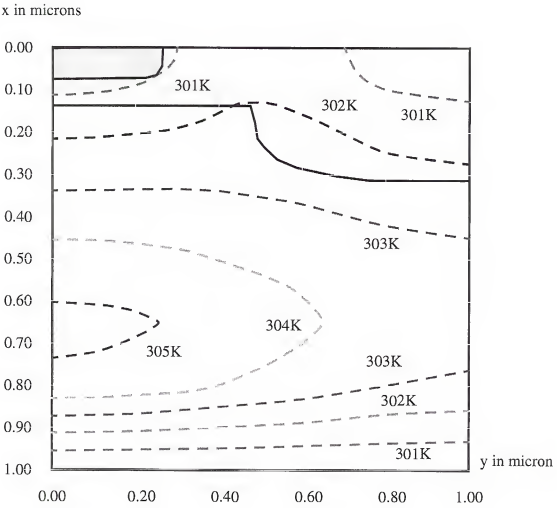


Figure 2.3. The lattice temperature contour in an npn BJT at the biases of Vbe = 1.0V and Vbc = 3V. The ambient temperature is 300K and the thermal conductances at the contacts are infinite. Dark-shaded area is the emitter and light-shaded area is the base.

layer where the hot-carrier effect takes place due to the high electric field. The electron temperature in the region is as high as 4500K and the energy transfer from electron to other subsystems is significant. However, the peak lattice temperature in the region is only 6K higher than the boundaries. The base and collector currents are mainly determined by the drift-diffusion in the emitter and base regions, respectively, and the self-heating effect shows up in the physical models such as intrinsic carrier density and mobility. Since the temperature distribution is uniform in these regions, the constant temperature assumption is valid in the BJT devices. The comparisons of the running time and error between these three simulations are summarized in Table 2.1. At the biases of V_{be}=0.9V and V_{be}=3V, the error of the SNSH model is less than 2% but the solution time is 20% less than the TD simulation, whereas the error of the isothermal simulation is greater than 7% and running time is slightly faster than the SNSH simulation.

Table 2.1 The comparisons among three simulation schemes for BJTs

Model	Vemit = -0.8V, Vcoll = 3.0V		Vemit = -0.9V, Vcoll = 3.0V	
	CPU time	error (%)	CPU time	error (%)
Full-thermo-dynamic	1	0	1	0
Single Node Self-Heating	0.82	0.10	0.77	1.52
Isothermal	0.78	1.17	0.71	7.09

The lattice temperature along the surface of a 0.6μm NMOS with TD simulation is plotted in figure 2.4 when the biases at drain and gate are 7V. Due to the thermal conductances at all contacts, the device is heated up and the lattice temperature at the boundaries is at least 40K higher than the ambient temperature. At the same operation condition, the lattice temperature of SNSH and isothermal simulations is 345K and 300K, respectively. Since the

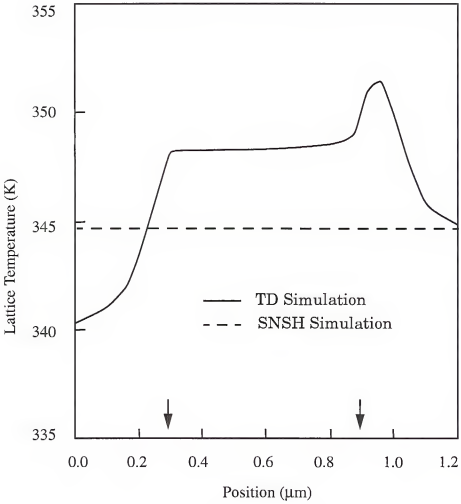


Figure 2.4. Lattice temperature along the surface of a NMOS from TD and SNSH simulations. Both the biases at gate and drain are 7V and the ambient temperature is 300K. A thermal conductance of 9000 K/Wcm is attached to each contact. The arrows mark the gate width.

drain current is governed by the electron transport in the channel and the self-heating effect shows up in the mobility model, the large difference of channel lattice temperature between isothermal and TD simulations results in significant error in drain current. In contrast, the discrepancy of drain current between SNSH and TD simulations is little due to the small difference of the channel lattice temperature. Figure 2.5 shows the comparisons among the TD, SNSH, and isothermal simulations of the drain current as a function of biases. The comparison between TD simulations and experimental data is shown in Liang and Law [Lia94b]. The drain current predicted by the TD model is always lower than that of the isothermal model because the channel lattice temperature in TD simulation is higher than that of the isothermal simulation which increases the lattice scattering and decreases the electron mobility. The degradation of the electron mobility predicted by TD model causes negative output conductance at high Vgs which is not observed in the isothermal simulation. The error of the isothermal simulation is larger than 10% when Vgs = Vds = 7V. For the SNSH simulation, the degradation of mobility and the negative conductance are also observed. The error of SNSH simulation is within 1.2% but the simulation time is 20% less than the TD simulation. The cpu time and error of the three simulations are summarized in Table 2.2.

The accuracy of this approach is determined by the difference between the lattice temperature of SNSH simulation and actual lattice temperature in the critical active regions, *i.e.* channel in MOSFETs and intrinsic base, emitter and base-collector space charge regions in BJTs. For a device with large temperature variations in a big critically active area such as a power transistor, the physical mechanisms can not correctly be described by a single lattice temperature. This could result in a larger error in the terminal current. However, this approach appears useful for a submicron device structure where

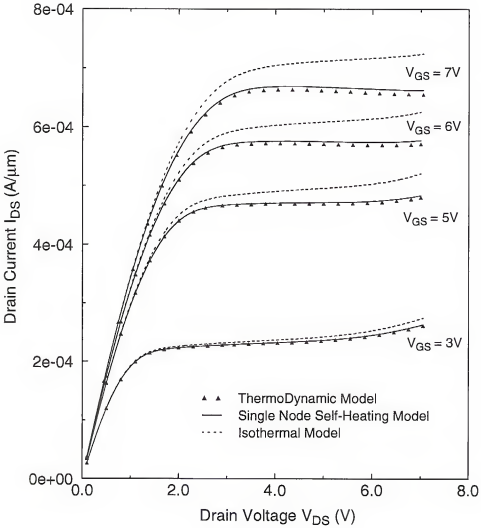


Figure 2.5. The comparisons of the I_{DS}-V_{DS} plot of a NMOS at different V_{GS}s between three models.

the lattice temperature variation in the critically active region is small and can be properly approximated by the SNSH simulations. The accuracy of the SNSH model is insensitive to the size of the noncritically active region since there is no heat generated in the area and it has little effect on the lattice temperature of the SNSH simulations.

Table 2.2 The comparisons among three simulation schemes for MOSFETs

Model	Vdrain = 7V, Vgate = 6V		Vdrain = 7V, Vgate = 7V	
	CPU time	error (%)	CPU time	error (%)
Full-thermo-dynamic	1	0	1	0
Single Node Self-Heating	0.80	1.02	0.72	1.16
Isothermal	0.76	8.69	0.78	10.5

2.3 Current Boundary Condition

In the analysis of BJT breakdown or CMOS latch-up triggering, an I-V curve snake back phenomenon is usually observed. The current is a multi-valued function of the applied voltage in that regime which implies that the numerical solutions of the analysis may end up with a solution in one of three distinct and stable states depending on the initial guess. Furthermore, in the negative resistance regime or around the snake back point where $dV/dI = 0$, it is very difficult or impossible to obtain a solution with a voltage boundary condition. Since the current is a single-valued function of the applied voltage in these regimes, the current boundary condition can provide a unique and faster converged solution than the voltage boundary condition.

For device simulators such as MEDICI, the boundary condition for the Poisson equation in the contact region is to fix the potential. The definition is invalid if there is carrier accumulation or depletion at the contact where the

relative position of the band-gap and majority carrier Fermi energy begins to change. In FLOODS, the voltage boundary condition for Poisson equation is defined by the fixed Fermi-level E_{FC} at the contact and given by [Lia94a]

$$E_{FC}(\psi, C, T_c) = E_{FCO}(\psi_o, C_o, T_{co}) - V_{applied} \tag{2-8}$$

where ψ, C and Tc represent the potential, majority carrier, and majority carrier temperature solutions, respectively. Subscript *o* indicates the equilibrium solution variables and $V_{applied}$ is the applied voltage at the contact. For the current boundary condition, eq. (2-8) is replaced by

$$J_{applied} = \sum_i (J_{ni}(\psi, n, T_n) + J_{pi}(\psi, p, T_p) + J_{dis,i}(\psi, n, p)) \tag{2-9}$$

where $J_{applied}$ is the applied current density at the contact, *i* indicates the contact nodes, *n* and *p* are electron and hole solutions, respectively, and J_{ni}, J_{pi}, and $J_{dis,i}$ are the electron, hole, and displacement current densities, respectively.

The dependence of current densities on potential ψ is weaker than on Fermi energy. Thus, in the nonlinear Newton iteration, the convergence of the current boundary condition is slower than that of the voltage boundary if the variation of ψ solution is large, *i.e.* dI/dV is small. However, the dependence of the current densities on carrier density is slightly stronger than the Fermi energy. Therefore, the convergence of the current boundary will be faster than that of the voltage boundary if the current ψ solution is very close to the previous one and the variation of the carrier solution is large in the Newton iteration, *i.e.* dI/dV is large. The Newton iteration steps required for two-dimensional BJT simulations with voltage and current boundary conditions

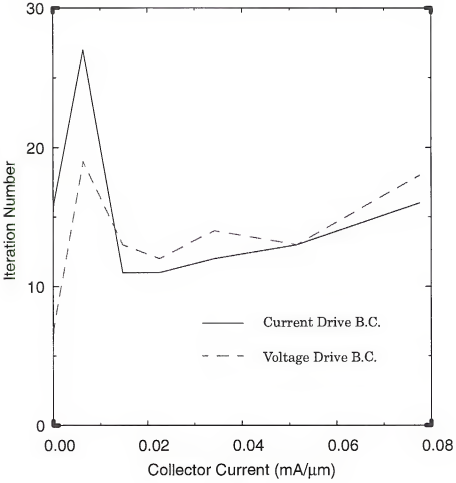


Figure 2.6. The number of Newton iteration required for normal active BJT simulations with current and voltage boundary conditions.

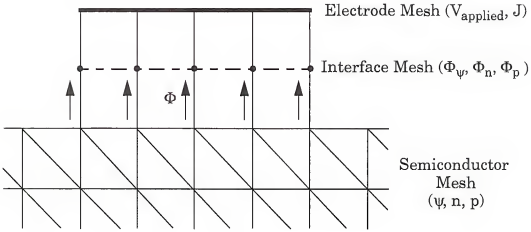
are shown in figure 2.6 as a function of collector current density. It can be seen that the voltage drive boundary condition is preferred at low current regime where dI/dV is small. In contrast, the current drive boundary has quicker convergence in the high current regime where dI/dV is large.

2.4 Interface Transport Model of Polysilicon Emitter Contacts

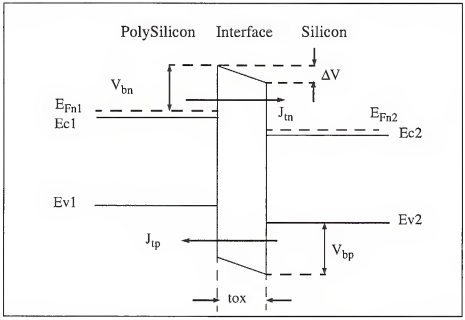
2.4.1 Motivation

By the use of heavily doped polysilicon as the diffusion source of emitters and emitter contacts, the fabrication of self-aligned, shallow junction bipolar transistors has been made possible and it leads to the features of higher current gain [Gra76, Gra79], packing densities, and much higher frequency capabilities [Suz85, Par86, Ich89].

A polysilicon layer is composed of mono-crystalline grains and disordered grain boundaries. Due to the potential barrier and recombination in the grain boundaries, polysilicon has a smaller carrier mobility and a shorter carrier life-time compared with mono-crystalline silicon [Set75]. A reduced mobility model in the polysilicon region has been proposed to explain the current gain enhancement in these devices based on the retarded transport mechanisms of minority carriers [Nin80]. During the fabrication process, a thin, oxide-like insulator is presented at the mono-crystalline silicon/polysilicon interface and this layer has a higher energy band-gap than silicon which acts as a potential barrier to carriers. Graaff [Gra79] proposed a carrier tunneling model to explain the carrier transport through the oxide-like layer and the current gain enhancement. These two mechanisms have been combined in the numerical models [Elt82, Yu84] which describe the experimental I-V characteristics more satisfactorily. Chor [Cho93] further applied the combination model to analyze the dependence of high frequency performance on the polysilicon



(a)



(b)

Figure 2.7. (a) The mesh structure in a contact. The boundary conditions are implemented inside the interface nodes and the solution variables of each node are the flux came from the corresponding semiconductor node. (b) The band diagram of a polyemitter contact.

emitter transistors and showed a significant difference in the charge storage behaviors with and without the presence of polyemitter contact.

Many parameters associated with the properties of the grain and grain boundary in the polysilicon material are required in the retarded mobility models. However, the physical properties of the polysilicon are not well understood and vary with different process technologies. Fossum [Fos80] proposed to characterize the minority-carrier behavior in polysilicon with an effective mobility and a lifetime constant, thus resulting in a reduced effective surface recombination velocity (ERV) as viewed from the monocrystalline emitter. The concept of ERV greatly simplifies the boundary condition used in the polyemitter contacts of BJTs in the device simulations. Though the use of constant ERV well characterizes the base current at low and intermediate bias, the current at high biases is not properly described due to the nonlinear I-V characteristics of the interface tunneling current [YU84]. Thus, the boundary condition with the combination of ERV and the current tunneling model is used in FLOODS to appropriately represent the properties of a polyemitter contact.

2.4.2 Model Implementation

To simplify the implementation, no additional material or mesh is created to adopt the current tunneling model. The model is directly applied to the boundary conditions with the current algorithm of FLOODS. Figure 2.7(a) illustrates the boundary contact structure and the solution variables associated with each mesh in FLOODS. The governing equations of the boundary conditions are solved in the interface mesh where the solution variables are the flux Φ of potential and carriers. Because the fixed potential boundary condition is invalid in the presence of charge accumulation or depletion in the contacts, FLOODS uses a fixed Fermi energy boundary condition and assumes the varia-

tion of the majority carrier Fermi energy E_F is corresponding to the change of potential with opposite sign, *i.e.*, $dE_F/d\psi = -1$. Thus, the boundary condition of Poisson equation at the polyemitter interface can be written as

$$\Phi_\psi + \frac{\varepsilon_{ox}}{q \cdot t_{ox}} \cdot (E_F + V_{applied}) = 0 \tag{2-10}$$

where q is the electronic charge, ε_{ox} is the permittivity and t_{ox} is the thickness of the oxide-like insulator, and Φ_ψ is the potential flux through the interface due to the charge accumulation or depletion at the boundaries of the interface.

In an n-type polysilicon contact, the tunneling current flux through the insulator is given by [Yu84]

$$J_{tn} = -\frac{Ae^{-b_n}}{q(d_n k)^2} \cdot \frac{\pi d_n k T_n}{\sin(\pi d_n k T_n)} \cdot (1 - e^{-d_n q \Delta V}) \tag{2-11}$$

and

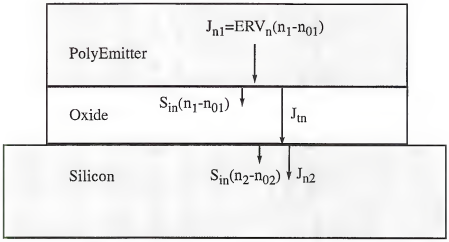
$$J_{tp} = \sqrt{\frac{kT_p}{2\pi m_p}} \cdot \frac{e^{-b_p}}{1 - d_p k T_p} \cdot \left(p_2 - p_1 \cdot e^{-\frac{q\Delta V}{kT_p}} \right) \tag{2-12}$$

where p_2 is the hole concentration at the monocrystalline side of the interface and p_1 at the polysilicon side, ΔV is the voltage across the interface which equals $E_{Fn2} + V_{applied}$ (refer to figure 2.7(b)), k is the Boltzmann constant, m_m and m_p are the effective masses of electron and hole, respectively, and A is the Richardson constant of electrons

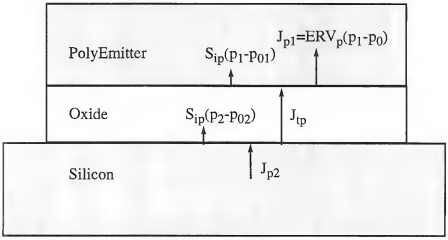
$$A = \frac{4\pi m_n q k^2}{h^3} \tag{2-13}$$

If ΔV is less than the barrier height of the interface, the coefficients b_c and d_c are given by

$$b_c = \frac{4\pi t_{ox}}{h} \cdot \sqrt{2m_{n,p} V_{b_{n,p}}} \tag{2-14}$$



(a)



(b)

Figure 2.8. The current components in each region of the polysilicon emitter contact for (a) electron and (b) hole. S_i is the surface recombination velocity of the interface, and n_0 and p_0 are the electron and hole densities at thermal equilibrium, respectively.

and

$$d_c = \frac{2\pi t_{ox}}{h} \cdot \sqrt{\frac{2m_{n,p}}{V_{b_{n,p}}}} \tag{2-15}$$

where h is the Planck constant, c represents electron n or hole p, and V_b is the average energy barrier height of the interface for electrons or holes.

The carrier flux components in each region of the polysilicon emitter contact are plotted in figure 2.8. The excess carriers ΔC_1 in the polysilicon side of the interface can be related to the carrier flux Φ_{C1} of the same region by $\Delta C_1 = C_1 - C_0 = \frac{\Phi_{c1}}{ERV_c}$. Then, the boundary condition for electron or holes in the interface mesh is given by

$$\left(1 \pm \frac{S_{ic}}{ERV_c}\right) \cdot \Phi_{c1} - J_{tc} = 0 \tag{2-16}$$

and the continuity equations for those nodes in the semiconductor mesh at the contact boundary can be expressed by

$$(J_{tc} \pm S_{ic}(C_2 - C_{02})) + \sum J_{c2} = 0 \tag{2-17}$$

where "+" for holes and "-" for electrons if $J_{tc} > 0$, and vice versa.

2.4.3 Results and Discussion

The comparison of the simulated Gummel plot of a polyemitter BJT with different thickness of interface layers is shown in figure 2.9. Without an interface layer, the base current remains a straight line at high biases and the collector current base push-out effect takes place at $V_{be} = 0.96V$. With the introduction of the interface layer, both the base and collector current deviate from the curves without the presence of the interface layer (ideal curves). The thicker the layer, the larger the current deviation.

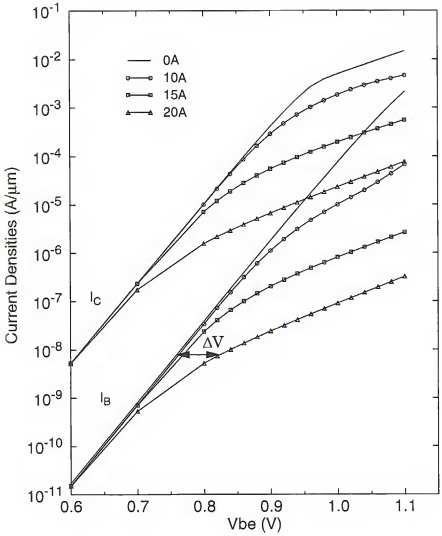


Figure 2.9. The simulated Gummel plot with different interface thicknesses. The voltage deviation $\Delta V = R_B I_B + R_E I_E$.

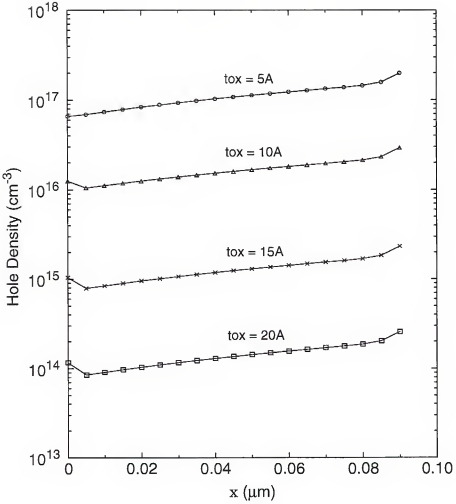


Figure 2.10. The hole densities in the emitter of a BJT with different thicknesses of interface layers. The BJT is biased at V_{be} = 1.0V and V_{cb} = 2.0V. The interface layer locates at x = 0.0μm and base-emitter junction is at x = 0.85μm.

There are two major effects associated with the presence of the insulator interface in the polyemitter contact. The first one is that the finite tunneling probability limits the carrier transport through the interface and results in an accumulation of holes and depletion of electrons at the mono-crystalline emitter side of the interface in an npn BJT. A thicker interface causes more hole accumulation and electron depletion due to the smaller tunneling probability. However, the collector current is governed by the characteristics of electrons in the base region and is not affected by the electron depletion in the emitter interface region. On the other hand, the base current is mainly the diffused hole current in the bulk region of a constant doped emitter. Figure 2.10 shows the hole densities in the emitter region with different interface thicknesses. It can be seen that the hole densities are lower in the entire emitter region for a thicker interface layer. Though there are holes accumulated in the interface side of the emitter, the hole densities determining the base current are dominated by the holes injected from the base-emitter junction. Thus, the carrier accumulation and depletion at the edge of the interface is not an important effect for the reduction of base and collector currents at high biases.

Another effect of the interface insulator layer is to introduce an additional voltage drop across the insulator due to the charge distribution at both sides of the layer. The voltage drop can be equivalent to an extra emitter resistance. Figure 2.11 shows the simulated emitter resistance as a function of V_{be} with an interface thickness of $10\mathring{A}$. Due to the larger carrier tunneling probability through the interface at higher biases, the emitter resistance decreases as V_{be} increases. The extra voltage drop in the interface causes the reduction of the voltage across the base-emitter junction. The total voltage reduction across the base-emitter junction ΔV can be obtained from the actual

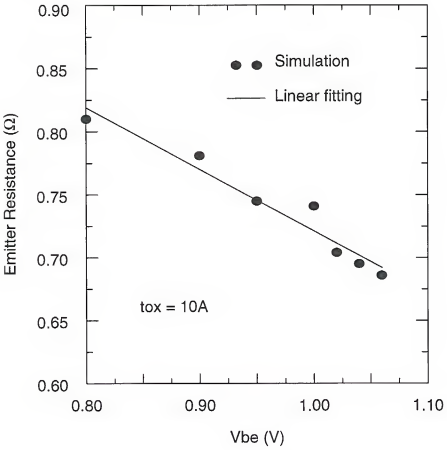


Figure 2.11. The resistance of the polyemitter interface from the FLOODS simulations. The thickness of the interface is $10\mathring{A}$ and the emitter area is $100\mu m^2$.

current deviation from the ideal base current in the Gummel-plot data and is given by

$$\Delta V = R_B \cdot I_B + R_E \cdot I_E \tag{2-18}$$

At low current injection, all the curves in figure 2.9 are close together because the reduction of the voltage across the base-emitter junction is not significant. When the current injection level becomes high, thicker interface with higher emitter resistance leads to a larger ΔV and a smaller voltage across the base-emitter junction. Thus, the hole density at the emitter side of base-emitter junction decreases following the increase of the interface thickness as can be seen in figure 2.10. Since ΔV has the same effect to base and collector current of a specific interface thickness, the base and collector currents begin to deviate from the ideal curves at the same bias.

2.5 Device/Circuit Tightly Coupled Simulator

2.5.1 Motivation

In high speed circuit design, circuit simulators such as SPICE are heavily used. One problem often faced is the inaccuracy of the active device models in the simulators which either require time intensive parameter extractions or nonphysical parameter fitting steps from measurement or device simulations. On the other hand, the device simulators such as PISCES provide more accurate predictions of a single active device with predetermined boundary conditions. However, the application of the device simulations to the circuit design is limited since it is difficult to predict the interaction between devices while the boundary conditions are not determined. The combination of device and circuit simulators is thus preferred in this respect.

For modern scaled devices, especially the high speed and high frequency devices, the resistances and capacitances in the active region are greatly reduced. However, the parasitics outside the active region are not scaled down accordingly and become important in high frequency operation. To properly predict the device performances, the inclusion of all the structures is necessary. Another advantage of the combination of device and circuit simulators is to cut down the number of grid points in the simulated device structure. Very often a structure contains purely resistive or nondistributed capacitive sections which can be represented by lumped resistances or capacitances. The replacement of these massive resistive or capacitive structures with lumped circuit components greatly reduces the number of grid points which in term saves cpu time and memory.

The coupled device/circuit simulators have been applied to the stead-state and transient simulations of BJT and MOS circuits [Rol88, Oka88]. However, the capability of a.c. simulations, one of the most important applications, is not available so far in this type of simulator. In this subsection, the implementation of the coupled device/circuit simulator and its applications to d.c. and a.c. simulations are presented.

2.5.2 Algorithms

There are two general methods to implement the device/circuit coupled simulator. "Loosely-coupled" approach [Oka88, May92, Kiz93] calculates the terminal informations (current I, voltage V) of the devices from the device simulator and the derivatives (dV/dI) numerically. The circuit simulator takes the informations from device simulations instead of an internal compact model to compute the node voltages and branch currents in the network. Several iterations between device and circuit simulators are required to get the

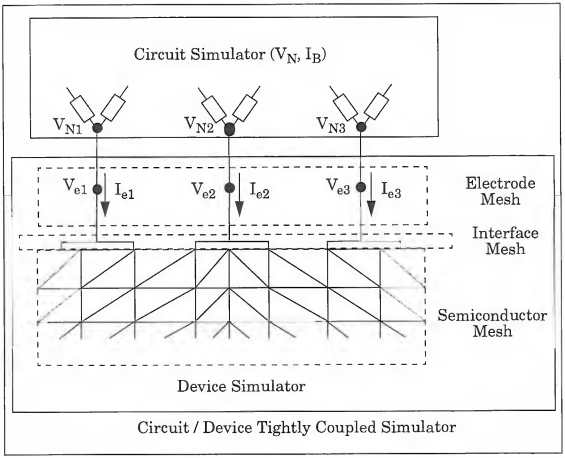


Figure 2.12. The diagram of the coupling algorithm between circuit and device simulators. The solution variables in device are ψ, n, p,., and the solution variables in circuit are V_N and I_B. The coupling is through the solution variables V_{ei} and I_{ei}, where i = 1, 2,...

$$
\begin{pmatrix}
S & & & \\
 & I & & \\
 & & E & \\
 & & & C
\end{pmatrix}
\begin{pmatrix}
\Delta\psi_i \\ \Delta n_i \\ \Delta p_i \\ \vdots \\ \Delta J_{\psi i} \\ \Delta J_{ni} \\ \Delta J_{pi} \\ \Delta V_{ei} \\ \Delta I_{ei} \\ \Delta V_N \\ \Delta I_B
\end{pmatrix}
=
\begin{pmatrix}
\text{P.E.} \\ \text{E.C.E.} \\ \text{H.C.E.} \\ \\ \text{eq. (2-19)} \\ \text{Elec. B.C.} \\ \text{Hole B.C.} \\ \text{eq. (2-20)} \\ \text{eq. (2-21)} \\ \text{M. N. E.}
\end{pmatrix}
$$

Figure 2.13. The Jacobian matrix of the coupled device/circuit simulator. The sub-matrices S and I are assembled from the semiconductor and interface boundary equations in the device simulator, respectively, and the sub-matrix C is from the modified nodal equations in circuit simulator. An additional sub-matrix E is assembled to couple the solutions between circuit and device.

solutions. This method is easy to implement but iterations between circuit and device simulations are time consuming and may cause divergence if the initial guess and the derivatives of the device terminal characteristics are not well estimated.

A "tightly-coupled" method [Rol88] solves circuit and device equations simultaneously using fully coupled and self-consistent algorithms. This method requires no iteration between device and circuit simulations and exhibits better convergence than loosely-coupled method. Due to the features of faster convergence and stability, the tightly-coupled algorithm is used in our device/circuit simulator.

The algorithm of the tightly-coupled device/circuit simulator is illustrated in figure 2.12. The circuit simulator is implemented with modified nodal equations (MNE) and the solution variables are the nodal voltages V_N and branch currents I_B of voltage sources. The basic structure of the device simulator is intact except the governing equations in the interface and electrode meshes. The boundary condition of Poisson equation (P.E.) in the interface is given by

$$E_F + V_{ei} = 0 \tag{2-19}$$

where V_{ei} is the contact voltage which is a solution variable in the electrode mesh and coupled to the circuit nodal voltage V_{Ni} by the equation in the electrode mesh

$$V_{ei} - V_{Ni} = 0 \tag{2-20}$$

The contact current I_{ei} is another solution variable in the electrode mesh which is also a branch current of circuit node Ni. The coupling of I_{ei} with device is given by the governing equation in the electrode mesh

$$I_{ei} = L_i \cdot (\sum J_{ni} + J_{pi} + J_{dis,i}) \tag{2-21}$$

where L_i is the contact length which converts the current density of 2D device simulator (in unit of A/μm) to current that flows through the circuit network if the uniform current densities at the contacts of device are assumed. The resulting Jacobian matrix of the coupled simulator and the governing equations used are plotted in figure 2.13. The sub-matrix S represents the semiconductor and I is the interface Jacobian from the device simulator and sub-matrix C is the Jacobian of circuit simulator. The link of these two systems is through the variables V_{ei} and I_{ei} in the electrode mesh of the device simulator. V_{ei} and I_{ei} serve as the terminal solutions and boundary conditions of the device as well as the initial values and solutions of the circuit nodes simultaneously. The strong coupling between these equations and solution variables ensures the fast convergence of the algorithm.

In FLOODS, the a.c. simulation scheme is implemented with the sinusoidal steady-state analysis techniques [Lau85]. The nodal equations of a capacitor C is $i_c = C \cdot \frac{d(v_h - v_t)}{dt}$ and inductor L is $v_h - v_t = L \cdot \frac{di_L}{dt}$, where v_h and v_t are the small-signal nodal voltages at each side of C or L, and i_c and i_L represent the small-signal branch currents of C and L, respectively. The Stiffness matrix expressions for the sinusoidal steady-state solutions of C and L are given respectively by

$$\begin{bmatrix} jwC & -jwC \\ -jwC & jwC \end{bmatrix} \cdot \begin{bmatrix} v_h \\ v_t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{2-22}$$

and

$$\begin{bmatrix} -\frac{j}{wL} & \frac{j}{wL} \\ \frac{j}{wL} & -\frac{j}{wL} \end{bmatrix} \cdot \begin{bmatrix} v_h \\ v_t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{2-23}$$

After assembling the global a.c. matrix, the a.c. system is

$$\left[J + jD\right] \cdot \tilde{X} = B \tag{2-22}$$

where J is the d.c. Jacobian matrix, D contains Stiffness matrices of time derivative terms such as those of C and L, $\tilde{X}$ is the a.c. solution vector, and B is the vector contains the a.c. voltage and current sources of the circuit. By splitting the system into real and imaginary part, eq. (2-22) can be rewritten using only real arithmetic as

$$\begin{bmatrix} J & -D \\ D & J \end{bmatrix} \cdot \begin{bmatrix} \tilde{X}_R \\ \tilde{X}_I \end{bmatrix} = \begin{bmatrix} B_R \\ B_I \end{bmatrix} \tag{2-23}$$

where subscript R indicates the real part and I is the imaginary part of the vector. The solutions of the small-signal analysis can be obtained by solving the matrix.

2.5.3 D.C. Simulations

A long-stripe, single-metal, double-poly-base-contacts BJT (see figure 2.14) is used for the studies. This device is designed by Harris Semiconductor Inc. for radio frequency applications. Usually, only the active region of the device, as shown in figure 2.15, is simulated. Conventional simulations assume the device structure is symmetric to the center of the emitter and the biases at both base poly contacts are the same. Thus, only half of the BJT structure is required in the simulations which can get the same solutions as the whole active region simulations but the cpu time is about 60% less. The validity of the assumption will be discussed in later sections.

Figure 2.16 shows the simulation diagram of the device with lumped parasitics, where R_{pc} represents the resistance between metal and poly base contacts and is determined from the measured poly sheet resistance, R_c is the

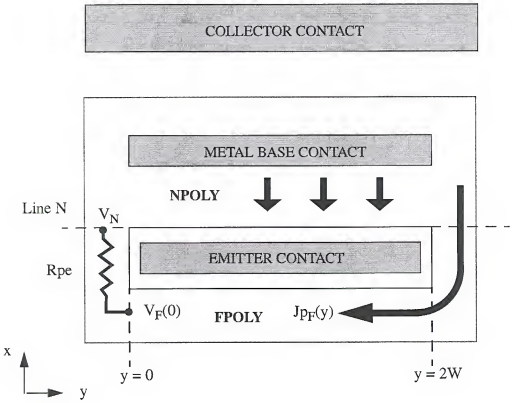


Figure 2.14. Top view of a long stripe double poly base BJT used in this work. The arrow signs indicate the low and high resistive current paths for NPOLY and FPOLY, respectively.

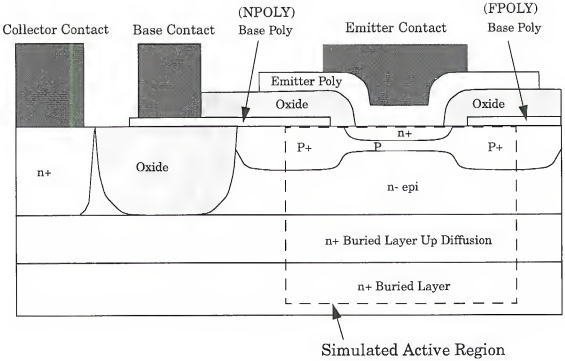


Figure 2.15. The cross section diagram of a single metal, double poly base contacts BJT. The area within the dashed line is the simulation region.

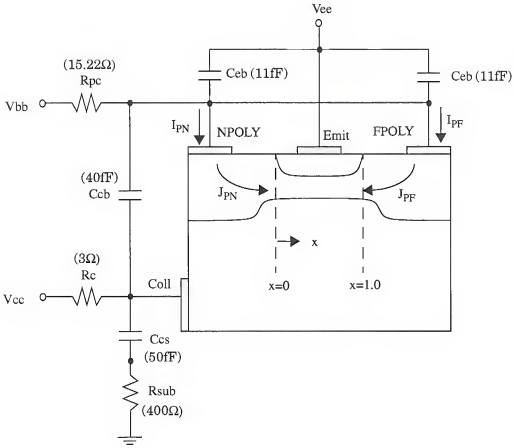


Figure 2.16. Schematic diagram of a BJT with parasitics used in the symmetric simulations. The NPOLY and FPOLY contacts are shorted. The total current flowing through each contact of the active device I_P equals the current density J_P times the actual contact length L_{actual}. *i.e.* $I_P = L_{actual} * J_P$. The parasitic values are for a device with 50μm emitter stripe.

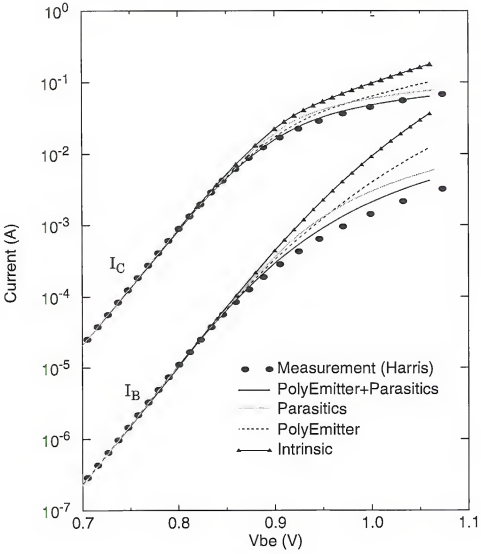


Figure 2.17. The Gummel-plot data from measurement and simulations when Vbc = 2V.

collector resistance in the buried layer as well as contact and estimated from additional simulations, C_{cb} takes into account the capacitance between base poly and collector layer, C_{eb} is the capacitance between base and emitter poly layers, and C_{cs} and R_{sub} are the series capacitance and resistance connecting to substrate. These capacitive parasitics are obtained from separate measurements and/or simulations. The connected resistance R_{pe} between NPOLY and FPOLY is neglected which results in the symmetric biases at the poly base contacts as assumed in the conventional simulations. The current density in each contact of the device is multiplied by the contact length (= $2W$) to represent the current flowing in the circuit network.

The comparison of Gummel-plot between different simulations and measured data is shown in figure 2.17. In intrinsic simulation, neither parasitics nor polyemitter interface model is included. The base current of the intrinsic simulation is basically a straight line until the current crowding effect in the intrinsic base region occurs. Though it properly predicts the current at low biases, intrinsic simulation overestimates both the base and collector currents at high level current injection. When taking the parasitics into account, the base and collector currents deviate from the intrinsic curves due to the additional voltage drop in R_{pc} which reduces the voltage across the base-emitter junction. The inclusion of parasitics improves the prediction but a large discrepancy is still observed in the simulation. If the polyemitter model with a $10\mathring{A}$ interface is also included, both base and collector currents further bend down due to the extra voltage drop in the interface and are in good agreement with the measured data. Thus, both the external base resistance and polyemitter interface are involved in the high current effect and need to be included in the simulations for accurately predicting the d.c. characteristics of devices.

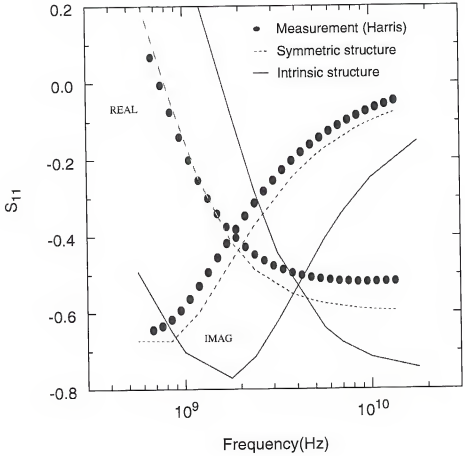


Figure 2.18. The s-parameters comparisons between Measured data, intrinsic device simulation, and simulation with parasitics (Symmetric structure). (a) input reflection coefficient S_{11}.

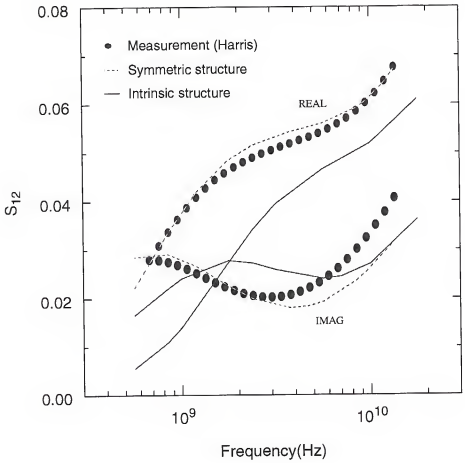


Figure 2.18. (b) Reverse transmission coefficient S_{12}.

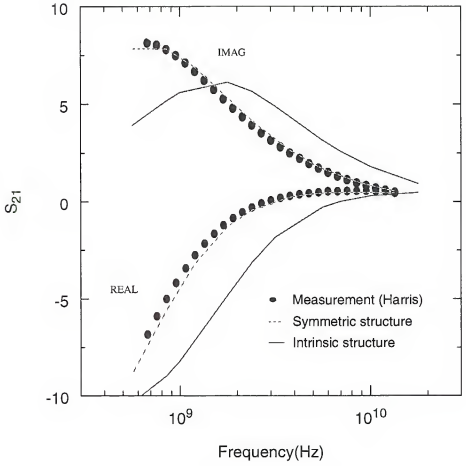


Figure 2.18. (c) Forward transmission coefficient S_{21}.

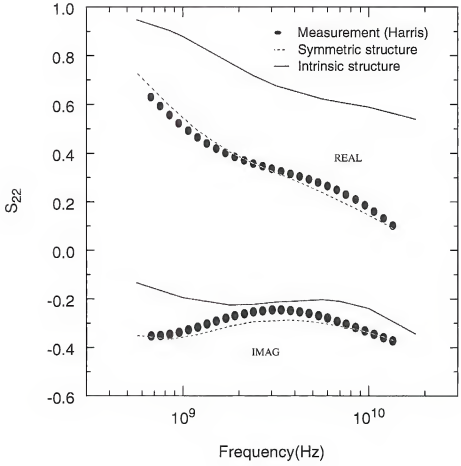


Figure 2.18. (d) Output reflection coefficient S_{22}.

2.5.4 A.C. Simulations

To characterize the high frequency performance, scattering parameters (s-parameters) are used because they are easier to measure at high frequencies than other kinds of parameters [Sze81]. S-parameters are measured using a matched termination. For example, to measure the S_{11} we measure the input reflection coefficient when the output port is terminated with an impedance equal to the characteristic impedance of the transmission line. Using the matched termination, a traveling wave incident on the load will be totally absorbed and no energy will be returned to the output port. Using the s-parameters to measure a transistor has an advantage that the transistor does not oscillate. In contrast, if we were to use a short- or open-circuit test, the transistor could become unstable [Gon84]. However, it is difficult to properly describe the transmission line with finite RLC circuit elements in the circuit simulations. Thus, the y-parameters are solved in our device/circuit simulations and then converted equivalently to s-parameters [Gon84].

The structure shown in figure 2.16 is used for the a.c. simulations. The capacitive parasitics C_{cb} and C_{cs} are adjusted to get the best fit of the measured data due to the difficult capacitance extraction. The resistive parasitics are intact in the simulations since the extraction is straight-forward. The simulated s-parameters of a common-emitter configuration BJT are plotted in figure 2.18 together with the measured data. It is obvious that none of the s-parameters is well predicted by the intrinsic simulations. The simulated maximum available power gain frequency f_{mag} is 19.5GHz which is much higher than 10.3GHz of the measurement. With the coupled device/circuit simulator, the parasitic effects can be taken into account and the predicted f_{mag} is 11.1GHz. The accuracy of the simulated s-parameters is significantly improved as can be seen in figure 2.18.

The sensitivity of the simulations to the parasitics is also analyzed. It is found that, among the parasitics, the base external resistance R_{pc} (NPOLY resistance) is the most important parameter which strongly affects all s-parameters. Since the input reflection coefficient S_{11} depends only on the input impedance of the device which is the series of the intrinsic base impedance and R_{pc}, thus, R_{pc} is the only parasitic controlling S_{11}. C_{cb} connects the base and collector electrodes which is important to the reverse transmission coefficient S_{12} at high frequencies due to the direct conduction between input and output ports. Though the effective capacitance of C_{cb} in the input port is amplified due to the Miller effect, C_{cb} has little effect on the forward transmission coefficient S_{21} and S_{11} since the Miller capacitance of C_{cb} is parallel with C_{π}, the sum of the diffusion capacitance in base-emitter junction and the Miller capacitance in base-collector junction, which predominates the input capacitance. Except C_{pe}, the output reflection coefficient S_{22} is sensitive to all the parasitics, especially the substrate parasitics C_{cs} and R_{sub} as well as the base resistance R_{pc}. Due to the parallel with the much larger capacitance C_{π}, C_{pe} shows no effect on any s-parameter.

The number of node in the simulated device mesh is 4348 which is much larger than the number required by the circuit simulator. Therefore, the increase of the running time in the coupled device/circuit simulations is indistinguishable compared with the intrinsic device simulations. Since the circuit elements are all linear in the simulations, the convergence rates also are the same in both simulations.

2.6 Summary

Several new models and functions have been implemented in FLOODS: The single node self-heating model has successfully reduced the full thermody-

namic simulation time while a good accuracy was achieved. The good convergence ability of the current boundary condition has also been demonstrated. The implementation of the polyemitter interface model has provided a better description and boundary condition for the polyemitter BJT and the improvement of the Gummel-plot data has been obtained. A circuit simulator has been developed and tightly coupled to FLOODS. By taking into account the proper interface and parasitic effects, good agreements have been achieved between simulations and d.c. as well as a.c. measurements.

CHAPTER 3
THREE DIMENSIONAL BASE DISTRIBUTED EFFECTS OF STRIPE BJTS--BASE RESISTANCE AT DC

3.1 Introduction

Base resistance is important for characterizing the large- and small-signal behavior [Tan79], switching characteristics [Spi65, Gau79] and noise performance [Zie58, Auf96] of high speed bipolar transistors. The accurate determination of the base resistance is essential for BJT design and modeling. It is known that the base resistance decreases as the operating current increases due to base conductivity modulation, base push-out in high-level injection, and current crowding effects [Yua88, Sat90,Fus92]. The crowding is the tendency for current to flow at the edge of the intrinsic base region at high current injection levels due to a voltage drop along the base current path in the intrinsic base region. The current crowding reduces the effective base current path and decreases the base resistance. Accurate modeling of base resistance is complicated due to its distributed nature and operating point dependence.

Several papers have discussed modeling and computation of the base resistance based on the debiasing and current crowding effect in the two-dimensional (2D) intrinsic base region of BJTs [Yua88, Rhe95, Gho65, Ver91, Cho90] with the assumption that the voltage and current distribution in the third dimension are uniform. Schroter [Sch91] proposed a purely geometry-dependent model to compute the external base resistance. However, in contemporary high speed double polysilicon BJTs, a long stripe (see Fig. 2.14) or long

interdigitated base structure is often used to obtain high current drive and low base resistance. In this situation similar debiasing effect occurs in the polysilicon region as in the intrinsic base region which causes nonuniform voltage and current distribution in the third dimension of the device. The voltage and current distribution characteristics vary at different base-emitter biases resulting in the modulation of base resistance. To date, the debiasing effect in the base polysilicon region and the three dimensional (3D) nonuniform current distribution in the base along with its relationship to the base resistance at different biases has not been studied.

The purposes of this chapter are to model the d.c. distributed effect in both the intrinsic and base polysilicon regions and to develop a method for computing the base resistance from the quasi-3D model. Using the results, effects of scaling and silicide technologies on the base resistance and device design consideration are also examined.

3.2 Asymmetric Distributed Model in the Active Base Region

3.2.1 Modeled device structure

Figure 2.14 shows a typical stripe BJT with single metal, double polysilicon base contact structure which will be used to study the 3D voltage and current distributed effect. The base current in the polysilicon region can be divided into two major components. One flows from the metal base contact and injects directly into the intrinsic base region at the NPOLY (Near POLY) side. The other component flows through the FPOLY (Far POLY) before it is injected into intrinsic base region. Although the base current density flowing from metal contact to the edge of the NPOLY (lineN in Fig. 2.14) might be different, the voltage along lineN can be assumed to be constant (= V_N) due to the small

resistance and negligible ohmic voltage drop difference in the short and wide current path. The base current flowing through the FPOLY experiences larger resistance due to a longer and narrower current path compared with current in the NPOLY, which results in a larger debiasing effect in the FPOLY. The voltage varies along the FPOLY and is expected to be lower than that of the NPOLY.

To simplify the problem, we assume the base current injects from the base polysilicon contact to the intrinsic base region boundary through a lumped extrinsic base resistance R_{bext} without any current loss (see Fig. 3.1). The base current in the intrinsic base region is also assumed to flow only in the x-direction and has no current component in the y-direction. If the stripe is long, the base current contributed from the short edges of the emitter is negligible. The base current in the FPOLY can then be assumed to flow only in y-direction and conduct from the NPOLY through a lumped resistance Rpe at the edges of polysilicon base handle as shown in Fig. 2.14. Thus, the device can be treated as the integration of numerous slices of 2D structures in the y-direction.

3.2.2 Intrinsic Model Development

Based on the assumptions and derivation of Hauser [Hau64], the governing equation of the base current in the intrinsic base region $Jb(x)$ may be written as a second-order differential equation:

$$\frac{d^2}{dx^2}Jb(x) + \frac{\rho_b}{V_O}Jb(x)\frac{d}{dx}Jb(x) = 0 \tag{3-1}$$

where ρ_b is the intrinsic base resistivity, $V_O = \frac{KT}{q}$ is the thermal voltage, and x is the coordinate in the intrinsic base region. The solution of this equation is

$$Jb(x) = A \cdot \tan\left(A \cdot \frac{\rho_b}{2V_O} \cdot B\left(1 - \frac{x}{B}\right)\right) \tag{3-2}$$

where A and B are arbitrary constants of integration. The constants can be evaluated from the boundary conditions of the device. There is a point l in the intrinsic base at which the potential is a minimum and the base current is zero, *i.e.* $Jb(l) = 0$. This gives $B = l$. Equation (3-2) may then written as

$$Jb(x) = \frac{2V_O}{\rho_b} \cdot Z \cdot \tan(Z \cdot (l - x)) \tag{3-3}$$

where

$$Z = \frac{\rho_b \cdot A}{2V_O} \tag{3-4}$$

The voltage in the intrinsic base is then given by

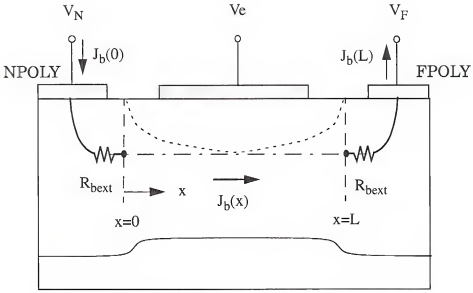
$$V(x) = V(0) - \int_0^x Jb(x) \cdot \rho_b dx = V(0) - 2V_O \ln\left(\frac{\cos(Z \cdot (l - x))}{\cos(Zl)}\right) \tag{3-5}$$

In case of a symmetric bias, i.e. V(0) = V(L), where L is the intrinsic base length beneath emitter, eq. (3-5) gives l = L/2. Substituting this into eq. (3-3), we obtain

$$Jb(L) = \frac{2V_O}{\rho_b} \cdot Z \cdot \tan\left(Z \cdot \frac{L}{2}\right) = -Jb(0) \tag{3-6}$$

Then Z is obtained from the solution of the equation.

$$\left(Z\frac{L}{2}\right) \cdot \tan\left(Z\frac{L}{2}\right) = \frac{\rho_b L}{4V_O} \cdot Jb(0) \tag{3.7}$$



— · — · — Path to extract potential in device simulation

- - - - - - - Base-emitter metallurgical junction in the device

Figure 3.1. Schematic diagram of the active region of a double polysilicon base BJT. Shaded regions are the polysilicon emitter and base. A lumped extrinsic base resistance Rb_{ext} connects the polysilicon contact and intrinsic base region. All the base current injected into the intrinsic base area flows through Rb_{ext}. The dashed line indicates the path used to extract potential from FLOODS simulations.

Combining eq. (3-3) and (3-5) at x=L, the relation between boundary current and voltage is given by

$$Jb(0) = \left(\frac{2V_O Z}{\rho_b \sin(ZL)}\right)\exp\left(\frac{V(0)-V(L)}{2V_O}\right) - \frac{2V_O Z}{\rho_b}\cot(ZL) \tag{3-8}$$

$$Jb(L) = -\left(\left(\frac{2V_O Z}{\rho_b \sin(ZL)}\right)\exp\left(\frac{V(L)-V(0)}{2V_O}\right) - \frac{2V_O Z}{\rho_b}\cot(ZL)\right) \tag{3-9}$$

If $|V(0) - V(L)| \ll 2V_O$, then

$$\exp\left(\frac{V(0)-V(H)}{2V_O}\right) \approx 1 + \frac{V(0)-V(H)}{2V_O} \tag{3-10}$$

The validity of this assumption will be verified in section 3.3.2. With this approximation, eq. (3-8) and (3-9) become

$$Jb(0) = \left(\frac{Z}{\rho_b \sin(ZL)}\right)(V(0) - V(L)) + \frac{2V_O Z}{\rho_b}\tan\left(\frac{ZL}{2}\right) \tag{3-11}$$

$$Jb(L) = -\left(\left(\frac{Z}{\rho_b \sin(ZL)}\right)(V(L) - V(0)) + \frac{2V_O Z}{\rho_b}\tan\left(\frac{ZL}{2}\right)\right) \tag{3-12}$$

and the total current is

$$Jbt = \frac{4V_O Z}{\rho_b} \cdot \tan\left(\frac{ZL}{2}\right) \tag{3-13}$$

3.2.3 Model with Extrinsic Base Resistance

Assuming all the base current in the intrinsic base region boundary flows through a lumped extrinsic base resistance Rb_{ext} from the polysilicon

base contact without any loss, as shown in Fig. 3.1, the I-V relation is then given by

$$V_N = V(0) + Jb(0) \cdot Rb_{ext} \tag{3-14}$$

$$V_F = V(L) - Jb(L) \cdot Rb_{ext} \tag{3-15}$$

where V_N and V_F are voltages at NPOLY and FPOLY, respectively. Substituting eq. (3-11) to (3-13) into (3-14) and (3-15), gives the terminal current density equations as a function of terminal voltage $V_F(y)$.

$$Jb_N(y) = Jb(0) = C_1 \cdot V_F(y) + C_2 \tag{3-16}$$

$$Jb_F(y) = -Jb(L) = -C_1 \cdot V_F(y) + C_3 \tag{3-17}$$

where y is the coordinate along the polysilicon base region,

$$C_1 = -\frac{\dfrac{Z}{\rho_b \cdot \sin(ZL)}}{1 + 2Rb_{ext} \cdot \left(\dfrac{Z}{\rho_b \cdot \sin(ZL)}\right)} \tag{3-18}$$

$$C_2 = \frac{Rb_{ext}Jbt \cdot \left(\dfrac{Z}{\rho_b \cdot \sin(ZL)}\right) + \dfrac{2V_O Z}{\rho_b} \cdot \tan\left(\dfrac{ZL}{2}\right) + \dfrac{Z}{\rho_b \cdot \sin(ZL)} \cdot V_N}{1 + 2Rb_{ext} \cdot \left(\dfrac{Z}{\rho_b \cdot \sin(ZL)}\right)} \tag{3-19}$$

and

$$C_3 = \frac{Rb_{ext}Jbt \cdot \left(\dfrac{Z}{\rho_b \cdot \sin(ZL)}\right) + \dfrac{2V_O Z}{\rho_b} \cdot \tan\left(\dfrac{ZL}{2}\right) - \dfrac{Z}{\rho_b \cdot \sin(ZL)} \cdot V_N}{1 + 2Rb_{ext} \cdot \left(\dfrac{Z}{\rho_b \cdot \sin(ZL)}\right)} \tag{3-20}$$

Though the base current density flowing from metal contact to the edge of NPOLY (lineN in Fig. 2.14) might be different, the voltage along lineN is assumed to be constant (= V_N) due to the small resistance and ohmic voltage drop in the short and wide current path. So if we assume ρ_b and Rb_{ext} are independent of current crowding at each bias and the total base current Jbt remains unchanged, then C_1, C_2, and C_3 become constants. Thus, the base current densities injected from NPOLY $Jb_N(y)$ and FPOLY $Jb_F(y)$ into the active base region are a linear function of $V_F(y)$, which is determined by the debiasing effect along the FPOLY.

.3.2.4 Verification and Discussion

A real single metal base contact (SMB) npn BJT with an emitter stripe length of 50μm is used for this study. The layout and cross-section of the intrinsic device structures are illustrated in Fig. 2.14 and 2.15, respectively. The 2D doping profile was determined using the process simulator Athena, and the polysilicon sheet resistance was obtained from independent measurements. The other resistances and boundary condition $J_b(0)$ required for the model are determined from 2D device simulations using FLOODS [15], [16]. Due to the Early effect and base conductivity modulation, the base width W_b and intrinsic base resistivity ρ_b vary depending on the current injection level and need to be extracted from simulations at each bias. The required parameters for the model of the transistor are summarized in Table 3.1.

Figure 3.2 compares the normalized voltage distribution along the intrinsic base region between the model and FLOODS simulations. In all cases, V_N is fixed at 0.8V and V_F is swept from 0.8V to 0.7994V. Both the model and simulations show similar debiasing effects across the intrinsic base region. The discrepancy between simulation results and the model is due to the fact that

Table 3.1 The device parameters required for model

parameters	value
Emitter width (Intrinsic base length), L	1μm
Intrinsic base width, W_b @Vbe=0.6V, Vcb=2V	0.22μm
Unit length extrinsic base resistance, R_{bext}	705.3Ω.μm
Intrinsic base resistivity, ρ_b @Vbe=0.6V, Vcb=2V	1.2KΩ.μm
Polysilicon sheet resistivity, ρ_p	16.7Ω.μm
Polysilicon thickness	0.288μm
FPOLY width	6μm
FPOLY cross section area, Ap	1.73μm^2
FPOLY connecting resistor, Rpe	12.57Ω
Resistance from lineN to metal, Rpc	13.74Ω
Emitter stripe length, 2W	50μm

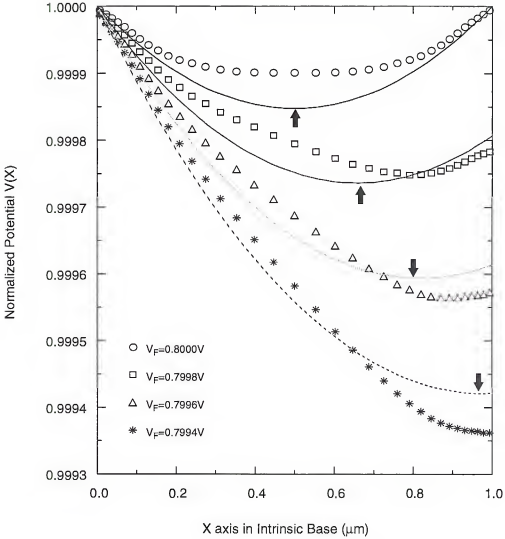
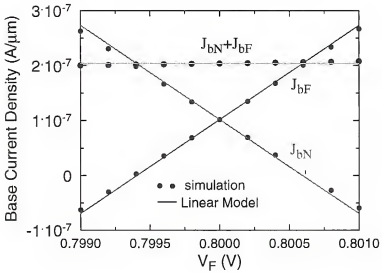


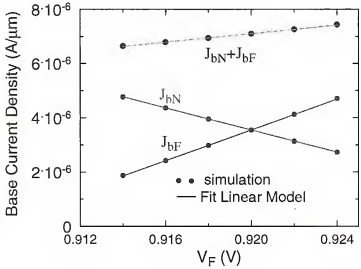
Figure 3.2. The DC potential profile in the intrinsic base region at different V_Fs (X=1.0 μm). In all cases, V_Ns (X=0μm) are fixed at 0.8V. The symbols indicate the device simulation data and lines are the modeling results.

the path used to extract the potential profiles in the simulations, as indicated by the dashed line beneath the emitter in Fig. 3.1, does not correspond to the actual base current path assumed in the model. Note, at the minimum voltage point l of the potential profile, where $J_b(x)$ is zero, the base current injected from base to emitter on the left hand side of l is contributed by the left terminal contact (NPOLY), and the current injected on the right hand side of l comes from the right terminal contact (FPOLY). The point l acts as a current divider in the intrinsic base region. When the bias is symmetric, l is located at the center of the intrinsic base and the current is equally divided between NPOLY and FPOLY. As V_F decreases, l moves toward FPOLY and forces part of the current originally flowing through the FPOLY at symmetric bias to flow through the NPOLY. Consequently, the current J_{bN} from the NPOLY increases corresponding to the decrease of current J_{bN} through the FPOLY.

The leverage behaviors of terminal currents at asymmetric biases can be observed from Fig. 3.3. The FPOLY current J_{bF} decreases and the NPOLY current J_{bN} increases if V_F is smaller than V_N. In the case of V_N=0.8V, the base current injected from base to emitter is approximately constant everywhere in the intrinsic base region since the voltage variation in the FPOLY and in the intrinsic base region are small. Thus, the reduction of J_{bF} is compensated by the increase of J_{bN}, and the total current density is approximately constant at different V_Fs as shown in Fig. 3.3(a). As the bias increases, a more significant ohmic voltage drop takes place along the FPOLY which reduces the voltage across the base-emitter junction and decreases the total base current density. Figure 3.3(b) shows the J-V_F relation while V_N is biased at 0.92V. The total current reduces by 6.7% when V_F is 6mV lower than V_N. Under these high injection conditions, the coefficient C_1 in Eq. (3-16) is slightly different from C_1 in Eq. (3-17). For better accuracy in the quasi-3D model, C_1s in Eq. (3-16) and



(a)



(b)

Figure 3.3. Simulated and modeled polysilicon base terminal current as a function of V_F at (a) V_N= 0.8V, and (b) V_N= 0.92V. In (b), C_1 in (3.16) and (3.17) is adjusted to meet the decreasing total current.

(3-17) are corrected according to the simulated slopes of J_{bN} and J_{bF} at high biases as illustrated in Fig. 3.3. Though it requires a correction in the coefficients, the linear relation between the terminal currents and V_F, *i.e.* Eq.(3-16) and (3-17), is valid until the voltage difference approaches $2V_o$ which is at a bias exceeding the range of interest.

3.3 Quasi Three Dimensional DC Distributed Model

3.3.1 Model Development

The following assumptions have been made while deriving the expressions for current and voltage distribution in the polysilicon base regions: (i) the resistivity of polysilicon base ρ_p is constant, (ii) the current density in the FPOLY is uniformly distributed in the direction perpendicular to the current flow (y-direction), and (iii) the device is symmetric in y-direction.

Let $Jp_F(y)$ be the current density of transverse base current in the FPOLY and be related to the terminal current $Jb_F(y)$ by the equation

$$Jp_F(y) = Jp_F(W) + \int_y^W \frac{Jb_F(y)}{Ap} dy \tag{3-21}$$

where Ap is the cross section area of the polysilicon base, and W is half of the emitter stripe length. Since the device is symmetric to the y-direction, we have $Jp_F(W) = 0$. The voltage in the polysilicon base $V_F(y)$ is given by

$$V_F(y) = V_F(0) - \int_0^y Jp_F(y)\rho_p dy \tag{3-22}$$

Substituting eq. (3-17) and (3-22) into (3-21) and differentiating twice, the result is

$$\frac{d^2}{dy^2} Jp_F(y) - C_4 Jp_F(y) = 0 \tag{3-23}$$

where

$$C_4 = \sqrt{\left(-\frac{C_1 \cdot \rho_p}{Ap}\right)} \tag{3-24}$$

Applying the boundary condition $Jp_F(W) = 0$ to the general solution of eq. (3-23), we have

$$Jp_F(y) = C_5 \cdot \sinh(C_4 \cdot (y - W)) \tag{3-25}$$

where C_5 is a constant and can be expressed as

$$C_5 = -\frac{Jp_F(0)}{\sinh(C_4 W)} \tag{3-26}$$

If eq. (3-25) is used in (3-22) and the integration performed, one obtains

$$V_F(y) = C_6 - C_7 \cdot \cosh(C_4 \cdot (y - W)) \tag{3-27}$$

where

$$C_7 = \frac{C_5 \cdot \rho_p}{C_4} \tag{3-28}$$

and

$$C_6 = V_F(0) + C_7 \cosh(C_4 \cdot W) \tag{3-29}$$

If the 3D nonuniform current flow at the short edge of the emitter stripe is neglected, and we assume all the current in the FPOLY flows through the resistance *Rpe* at the edge of base polysilicon, the edge voltage $V_F(0)$ of the FPOLY and V_N are related by

$$V_F(0) = V_N - (Ap \cdot Rpe) \cdot Jp_F(0) \tag{3-30}$$

Substituting eq. (3-26), (3-27) and (3-30) into the combination of (3-17) and (3-21), and the FPOLY boundary current is given by

$$Jp_F(0) = \frac{W}{Ap(C_8 - C_1 WRpe)} \cdot (C_3 - C_1 V_N) \tag{3-31}$$

which is only a function of V_N where

$$C_8 = 1 + \frac{C_1 \rho_p}{ApC_4^2} - \frac{C_1 \rho_p W}{ApC_4} \coth(C_4 W). \tag{3-32}$$

Once $Jp_F(0)$ is obtained, the voltage $V_F(y)$ and base current density $Jp_F(y)$ in the FPOLY can be solved, which in turn gives the voltage and current distribution in the whole area of the intrinsic base region.

3.3.2 Results and Discussion

The modeled ohmic voltage drop along the FPOLY at three different biases is shown in Fig.3.4. At Vbe=0.8V, the model shows a small debiasing effect along the FPOLY and the voltage variation is less than 0.3mV. When Vbe=1V, the ohmic voltage drop in the FPOLY is larger than 18mV due to the larger current flow in the FPOLY. However, even at this condition, $\Delta V \ll 2V_{TH} \approx 52mV$, the linear approximation in deriving Eq.(3-11) and (3-12) is still valid. The emitter current density injected through the base-emitter junction may change dramatically in the presence of the debiasing effect because of its exponential dependence on the base-emitter junction voltage. Figure 3.5 shows the modeled emitter current density *Je* beneath the emitter area when Vbe = 1.0V. *Je* is constant along the intrinsic base edge near NPOLY because of the constant bias at the NPOLY whereas *Je* decreases toward the center of the stripe due to the ohmic voltage drop along the FPOLY in the y-direction and in the intrinsic base region along the x-direction. Thus, the emitter current crowds at the edges of the stripe. The total emitter current

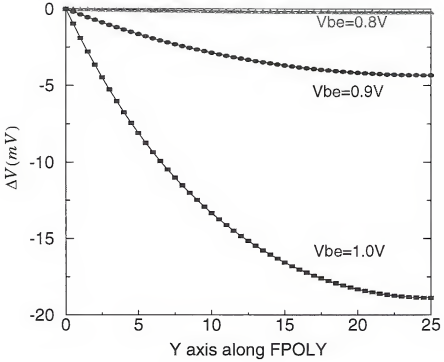


Figure 3.4. Modeled DC voltage variation in FPOLY at three different biases. The emitter stripe length is 50μm and the width of FPOLY is 6μm with an average sheet resistance of 58Ω/sq. The voltage is assumed to be symmetric around the center of FPOLY.

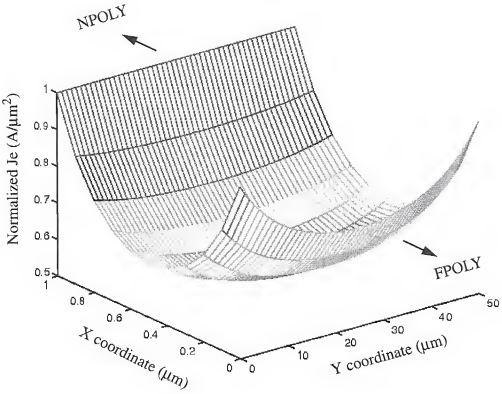


Figure 3.5. Plot of the two-dimensional distribution of emitter current calculated by the model. The intrinsic base sheet resistance is 4.1KΩ/sq.

computed from the quasi-3D model is smaller than that of the 2D model [Hau64] due to the additional debiasing effect in the y-direction. The base current density is proportional to the potential gradient in the intrinsic base region of xy-plane if the conductivity is constant.

Although the debiasing in the FPOLY at low biases is very small, it still cause the voltage in the FPOLY to be slightly lower than that in the NPOLY. The modeled terminal base current densities injected from the polysilicon to the active base region along the y-direction are shown in Fig. 3.6. The current, J_{bN}, injected from the NPOLY is always larger than the current, J_{bF}, from the FPOLY since the bias at the NPOLY is higher than the bias at the FPOLY. The difference between J_{bN} and J_{bF} is larger at the center than at the edge of the stripe due to the debiasing along the FPOLY. Note that the decreasing current in the FPOLY is compensated by the increasing current in the NPOLY if V_N is less than 0.9V. The total current is about the same as that without considering the debiasing effect in the FPOLY. However, when V_N is larger than 0.9V, the current decrease in the FPOLY is larger than the increase in the NPOLY because of the significant ohmic voltage drop in the FPOLY. The total current is less than that without taking into account the debiasing effect in the FPOLY.

3.3.3 Model Verification/Application

For a symmetric double polysilicon, double metal base contact structure, the base current contributed from each side of the polysilicon base should be the same. However, for the single metal base contact device as illustrated in Fig. 2.14, the current contributed by the FPOLY is less than that from the NPOLY due to the asymmetric bias as has been discussed above. Figure 3.7 shows the ratio of base current contributed from the FPOLY and NPOLY as a function of bias. Note that even at low biases, the ratio is less than 40% and

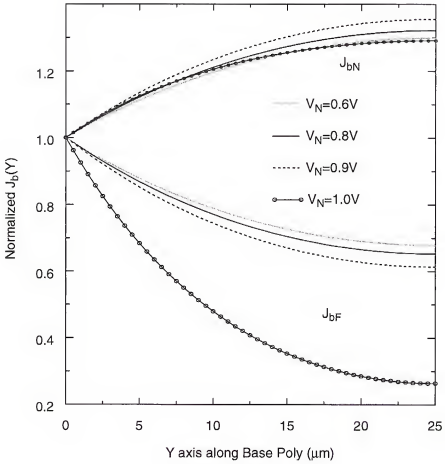


Figure 3.6. Modeled DC base terminal current density injected into active base region vs. position along base polysilicon at different biases. The stripe length is 50μm and voltage in the NPOLY is assumed to be uniform. The current densities J_{bN} and J_{bF} are normalized to that at y=0.0μm.

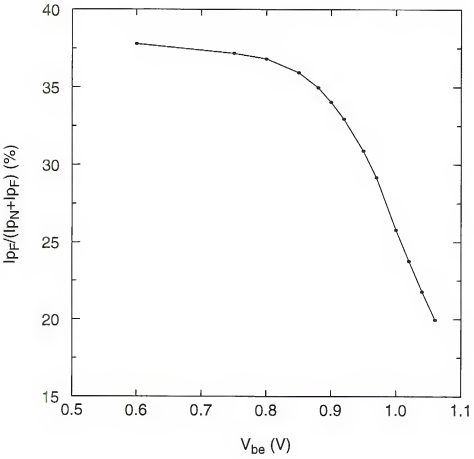


Figure 3.7. The ratio of base current contributed by FPOLY, I_{PF}, to the total current, $I_{PF}+I_{PN}$, as a function of bias calculated by the quasi-3D model, where I_{PN} is the current contributed by the NPOLY.

falls off slightly as the bias increases. For biases larger than 0.9V, the ratio decreases very fast due to the significant debiasing effect in the FPOLY.

The current flowing through the FPOLY experiences a larger base resistance than that through the NPOLY. Since the current via FPOLY decreases and the current associated with NPOLY increases at higher biases, it implies more and more current flows through the low resistive path as the bias increases. Thus, both the current crowding effect in the intrinsic base region and the current redistribution in the polysilicon base region contribute to the decrease of base resistance as the bias increases.

The total base resistance R_b can be calculated using the dissipated power method [Dub94]

$$P_b = I_b^2 R_b = \sum_i \int_{v_i} J_b \bullet \nabla V_b \cdot dv_i \tag{3-33}$$

where v_i is the volume of the region of interest, I_b is the total base current, J_b is the current density flowing through v_i, and ∇V_b is the gradient of potential in the direction of current flow. Figure 3.8 shows the division of regions used to compute the individual power dissipation in each region. The total power dissipation, P_b is

$$P_b = P_{BIF} + P_{BIN} + P_{PF} + P_{BEN} + P_{BEF} + P_{PC} + P_{PR} \tag{3-34}$$

where $P_{PC} = I_b^2 R_{PC}$ is the power dissipated in NPOLY through a lumped resistance R_{PC}. $P_{pe} = 2R_{pe} \cdot (ApJ_{pF}(0))^2$ is the power loss in the polysilicon connection between NPOLY and FPOLY. P_{BEN} and P_{BEF} are dissipated power in the extrinsic base regions due to lumped resistances, R_{bext}, connected to NPOLY and FPOLY, respectively. P_{PF} is the power consumption in the FPOLY, P_{BIF} and P_{BIN} are the power loss in the intrinsic base regions due to the current flow associated with FPOLY and NPOLY, respectively. Since the base current

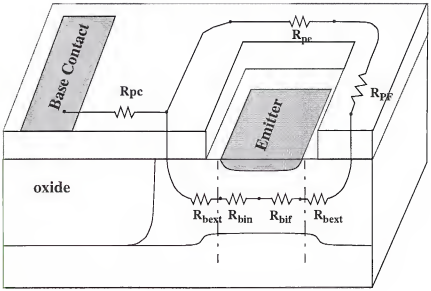


Figure 3.8. The diagram of the long stripe BJT structure and the definition of the base resistance in each subregion. The lightly shaded area is the base polysilicon region. Rpc, Rpe, and R_{bext} are the constant lumped resistances. R_{bin}, R_{bif} and R_{PF} are functions of the base current.

and voltage in each region have been obtained from the model, the total dissipated power and base resistance can be solved easily.

The measured and modeled large-signal base resistances are plotted in Fig. 3.9 as a function of bias. The Gummel-plot method [Fus95] is used to measure the DC base resistance. The 2D model curve indicates the resistance calculated with NPOLY and FPOLY shorted together which results in the symmetric voltage and current in both sides of the emitter stripe. Under this condition, $P_{BEN}=P_{BEF}$, $P_{BIN}=P_{BIF}$, and $P_{PF} = 0$. The base resistance of the quasi-3D model is seen to be much larger and more accurate than that of the 2D model in the bias range of interest. The analysis of the power dissipation in each region shows the difference comes from two facts. (i)the current flowing through NPOLY in a quasi-3D model has longer current paths and higher effective resistance than that of a 2D model due to the movement of the minimum voltage point toward FPOLY. (ii) the highly resistive current path of the FPOLY in the quasi-3D model.

As the bias increases, the base resistance of the 2D model mainly decreases due to current crowding and conductivity modulation in the intrinsic base region. In addition to the these mechanisms, the current redistribution from high to low resistive paths in the polysilicon region further reduces the base resistance in the quasi-3D model as the bias increases. The variation of base resistance computed by the quasi-3D model is thus larger and more accurate than that of the 2D model as can be seen in Fig. 3.9.

The 2D model represents a symmetric double metal base contact device structure (DMB) and quasi-3D model describes a single metal base contact structure (SMB). Figure 3.10 shows the modeled base resistance plotted as a function of emitter stripe ratio $\Re$ with three different FPOLY widths, λ, while V_{be} is biased at 0.9V. As expected, the base resistance of DMB decreases

inversely proportional to the increase of stripe length. The R_b of SMB always has higher values than that of DMB and the device with a narrower λ results in a larger R_b. The difference of R_b between these two structures is also included in Fig. 3.10. It is shown that the difference is small for the shorter stripe device and becomes larger as the stripe length increases because the distributed effects are more significant at a longer FPOLY. Scaling of the FPOLY width further enlarges the difference. For example, if $\Re$ is 10, R_b of SMB with λ=1μm is 26% higher whereas R_b of SMB with λ=2μm is 23% higher than R_b of DMB.

Silicided polysilicon technologies are often utilized in contemporary high speed double polysilicon BJTs to reduce the base resistance. However, this process increases the complexity of process and cost. Figure 3.11 shows the difference of R_b between DMB and SMB structures computed with different polysilicon sheet resistances and $\Re$. For a nonsilicided polysilicon with a sheet resistance of 100Ω/□ , R_b of SMB with $\Re$=6 is 21% higher than R_b of DMB. The difference increases to 36% if $\Re$=14. If the sheet resistance is reduced to 5Ω/□ , the difference decreases but is still significant for large $\Re$. Therefore, the design of SMB structure using symmetric device simulations without taking into account the distributed effects can significantly underestimate the base resistance. The results also suggest the R_b of DMB device is always much smaller than SMB structure unless a low sheet resistance polysilicon is used. The silicided technology is thus very important for the performance of SMB structure.

3.4 Summary

A quasi-3D model describing the distributed base voltage and current in both the intrinsic and polysilicon base regions has been derived. The results

illustrate the nonuniform characteristics of voltage and current in a long stripe transistor that cannot be properly described by a 2D model. A current redistribution effect in the polysilicon base regions due to the asymmetric bias modulation has been presented. In addition to the current crowding effect in the intrinsic base, this current redistribution effect also plays a key role in determining the bias dependent base resistance. The base resistance has been computed from the solution of the model and the accuracy has been verified by the measured data. The same concept and methodology can be applied to analyze the distributed effects in the long interdigitated structures.

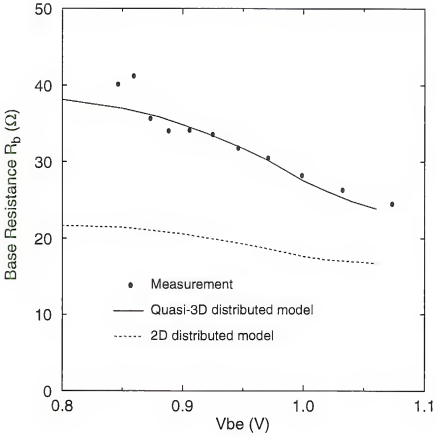


Figure 3.9. A comparison of DC base resistances obtained from measurements, 2D, and quasi-3D base distributed models. The emitter stripe length is 50μm and the width of the FPOLY in the device is 6μm.

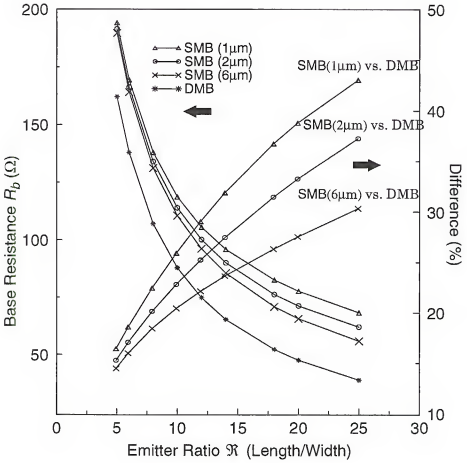


Figure 3.10. The plot of the modeled base resistance and difference between SMB and DMB structures as a function of emitter stripe ratio with different FPOLY widths. The sheet resistance of FPOLY is 58Ω/sq and FPOLY width is indicated in the parentheses.

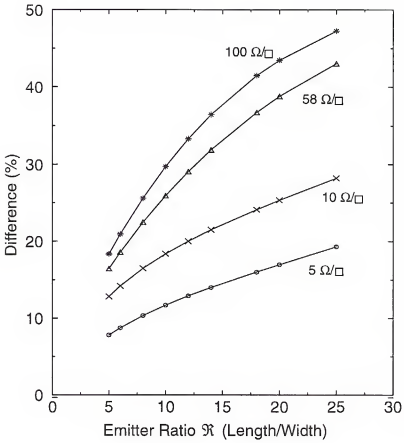


Figure 3.11. The difference of base resistance between SMB and DMB structures as a function of emitter stripe ratio with different FPOLY sheet resistances. The width of FPOLY is 1μm and V_{be}=0.9V.

CHAPTER 4
THREE DIMENSIONAL BASE DISTRIBUTED EFFECTS OF STRIPE BJTS--AC EFFECTS ON INPUT CHARACTERISTICS

4.1 Introduction

Design considerations for a microwave amplifier include stability, power gain, bandwidth, noise and d.c. bias requirements [Gon84]. A construction of an integrated circuit microwave amplifier usually starts with a proper design of the transistor; then the matching networks of the input and output ports are determined for a particular stability and gain criteria. An unconditionally stable transistor will not oscillate with any passive termination. On the other hand, a design using potentially unstable transistors requires careful analysis so that the matching network produces a stable amplifier. Thus, the accurate determination of the transistor characteristics is the first step to optimize the performance of a microwave amplifier.

The input impedance of a transistor in the common-emitter configuration is usually difficult to predict either by compact models or 2D device simulations due to the 3D distributed effect in the base region. This effect decreases the base impedance in magnitude with increasing frequency as a result of a high-frequency internal biasing effect. Several authors [Gho65, Pri61, Ver91] have modeled the base impedance with a uniform transmission line in the intrinsic base region to account for the a.c. distributed effect at high frequency operations. However, in double polysilicon bipolar transistors, the distributed characteristic in the polysilicon base region can also be significant

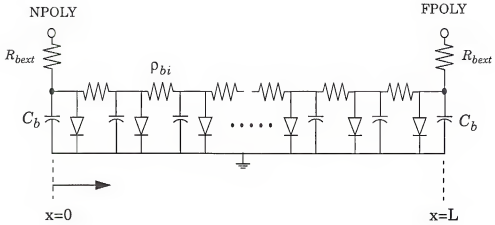
at high frequencies and no studies of this effect have been reported. Neglecting this effect may result in an inaccurate estimation of input impedance in the transistor and a poor design of an amplifier.

In this chapter, we concentrate on the distributed effect in the polysilicon base region, its effect on the input characteristics, and the accurate and efficient simulations of a stripe BJT at high frequencies. Since the distributed effect is a 3D problem, a quasi-3D simulation scheme is proposed to characterize the input behavior of the device. In the scheme, FLOODS is combined with an analytical model to deal with the voltage and current nonuniformity in the third dimension of the device due to the distributed effects in the base polysilicon region. The comparisons between quasi-3D, conventional 2D simulations and the measured data are performed and the results are discussed.

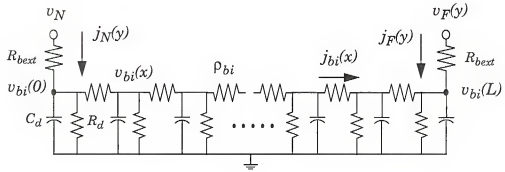
4.2 A.C. Distributed Model in the Active Base Region

4.2.1 Model Development

Figure 4.1(a) shows a distributed network representation of a rectangular active base region if the emitter and collector contacts are a.c. grounded [Gho65]. The diode represents the junction between base and emitter and the capacitance C_b accounts for the base-collector depletion capacitance under the emitter area and the geometric parasitic capacitance of the device. At a specific d.c. bias, the diode can be linearized by a parallel differential resistance R_d and a capacitance [Shr90] if the series inductance is negligible. The small-signal admittance of a section can be represented by $Y_d = R_d^{-1} + jwC_d$ at a specific angular frequency w, where C_d includes the diffusion and depletion capacitance of base-emitter junction and C_b. The complete small-signal network is plotted in figure 4.1(b).



(a)



(b)

Figure 4.1. Models of the distributed network in the active base region. (a) Large-signal network. (b) Small-signal network.

At high current densities, high-current effects like Kirk effect [Kir62] cause an increase of the transit time and hence a reduction of the cutoff frequency f_T. Because of this, we thus limit our interests in the current range where the maximum f_T is achieved. For the device used in this analysis, the maximum f_T is achieved at Vbe = 0.83V. Under this condition, d.c. debiasing effects in the intrinsic base as well as polysilicon base regions are not significant (see figure 3.4 and 3.6) and the current injection from base to emitter is approximately constant in the intrinsic base area. Thus, Y_d can be assumed to be uniform in the intrinsic base region, and the network becomes a uniform RC transmission line.

The equations governing the small-signal voltage $v_{bi}(x)$ and transverse current density $j_{bi}(x)$ in the intrinsic base region are given by

$$\frac{d}{dx}v_{bi}(x) = -\rho_{bi} \cdot j_{bi}(x) \tag{4-2}$$

$$\frac{d}{dx}j_{bi}(x) = -Y_d \cdot v_{bi}(x) \tag{4-3}$$

where ρ_{bi} is the average sheet resistance of the base layer. Substituting eq. (4-3) to (4-2), the resulting equation is

$$\frac{d^2}{dx^2}j_{bi}(x) = Y_d\rho_{bi}j_{bi}(x) \tag{4-4}$$

The solution of this equation is

$$j_{bi}(x) = C_1\sinh(\sqrt{Y_d\rho_{bi}}x) + C_2\cosh(\sqrt{Y_d\rho_{bi}}x) \tag{4-5}$$

Then,

$$v_{bi}(x) = -\sqrt{\frac{\rho_{bi}}{Y_d}}(C_1\cosh(\sqrt{Y_d\rho_{bi}}x) + C_2\sinh(\sqrt{Y_d\rho_{bi}}x)) \tag{4-6}$$

where C_1 and C_2 are two arbitrary constants to be determined from the boundary conditions. Solving $j_{bi}(x)$ and $v_{bi}(x)$ at x=0, it gives

$$C_1 = -\sqrt{\frac{Y_d}{\rho_{bi}}} v_{bi}(0) \tag{4-7}$$

and

$$C_2 = j_{bi}(0) \tag{4-8}$$

Substituting eq. (4-7) and (4-8) into (4-5) and (4-6) and solving at x=L gives the small-signal *j-v* relations

$$j_{bi}(0) = C_3 \cdot v_{bi}(0) - C_4 \cdot v_{bi}(L) \tag{4-9}$$

and

$$j_{bi}(L) = C_4 \cdot v_{bi}(0) - C_3 \cdot v_{bi}(L) \tag{4-10}$$

where

$$C_3 = \sqrt{\frac{Y_d}{\rho_{bi}}} \coth(\sqrt{Y_d \rho_{bi}} L) \tag{4-11}$$

and

$$C_4 = \sqrt{\frac{Y_d}{\rho_{bi}}} \operatorname{csch}(\sqrt{Y_d \rho_{bi}} L) \tag{4-12}$$

If the current densities at the edges of intrinsic base region flow through the lumped extrinsic base resistance Rb_{ext} without loss and equal to the current densities of poly contacts, then similar equations as (3-14) and (3-15) can be used to transform the terminal voltages at intrinsic base edges to that at base poly contacts. The results are:

$$j_N(y) = j_{bi}(0) = C_5 \cdot v_N - C_6 \cdot v_F(y) \tag{4-13}$$

$$j_F(y) = -j_{bi}(L) = C_5 \cdot v_F(y) - C_6 \cdot v_N \tag{4-14}$$

where

$$C_5 = \frac{C_3 + C_4 C_7 Rb_{ext}}{1 + Rb_{ext}(C_3 + C_4)} \tag{4-15}$$

$$C_6 = \frac{C_4 - C_4 C_7 Rb_{ext}}{1 + Rb_{ext}(C_3 + C_4)} \tag{4-16}$$

and

$$C_7 = \frac{C_3 - C_4}{1 + Rb_{ext} \cdot (C_3 - C_4)} \tag{4-17}$$

4.2.2 Verification and Discussion

The device structure shown in figure 2.14 and 2.15 is used for the a.c. analysis. The small-signal voltage v_N along lineN is assumed to be constant and the voltage along FPOLY $v_F(y)$ varies due to the debiasing effect. The small-signal admittance Y_d used in the model is obtained from the FLOODS simulations of two p-n junctions. One has the same doping profile as the modeled BJT in the base and emitter regions and is forward biased to extract the base-emitter junction admittance. The base-collector depletion capacitance is extracted from the reversed biased junction whose doping profile is the same as the modeled BJT in the base and collector region. The geometric parasitic capacitance is neglected in the evaluation of the admittance Y_d.

The normalized amplitude of the small-signal base current densities injecting from the intrinsic base region to emitter are plotted in figure 4.2 at different frequencies while the biases are symmetric. The current crowding effect is significant only if the oscillation frequency is higher than 1GHz. Figure

4.3 illustrates the same plot as figure 4.2 but with asymmetric biases. It can be clearly seen that the position of minimum voltage l moves to the lower bias terminal. The lower the frequency, the farther the point l moves. For the cases the operating frequencies are 1GHz and 100MHz, the point l is virtually out of the intrinsic base region. At this condition, all of the base current injecting from base to emitter is supplied by the higher bias terminal. Part of the current from the higher bias terminal flows to the lower bias terminal via the intrinsic base region. The current direction of the lower bias terminal is thus changed.

Figure 4.4 shows the small-signal base current densities at the poly terminals as a function of $v_F(y)$ oscillating at a frequency of 100MHz when v_N is biased at 3mV. If $v_F(y)$ is slightly smaller than v_N, the real part of the low bias terminal current density J_F decreases and the reduced amount of J_F is compensated by the increase of current density at the high bias terminal J_N due to the negligible decrease of total injection current. If $v_F(y)$ is much smaller than v_N, the current component flowing from high to low bias terminal is much larger than current injecting into emitter. The real part of the total small-signal current remains approximately constant as can be seen in figure 4.4(a). Since the applied biases $v_F(y)$ and v_N are real in this case, there is no imaginary current flowing from high to low bias terminal. The imaginary currents at both terminals decrease as $v_F(y)$ decreases which results in a decrease of the total imaginary current (figure 4.4(b)). Note the matched real current and the discrepancy of the imaginary current between the simulations and model. It implies the resistance is well modeled and the capacitance is not properly described by the uniform transmission line model. The reason is probably due to the assumption of the constant capacitance C_d in the model, which might not be true in the simulated device since the intrinsic base region is not rectangular and the doping density is not uniform along the base-emitter junction.

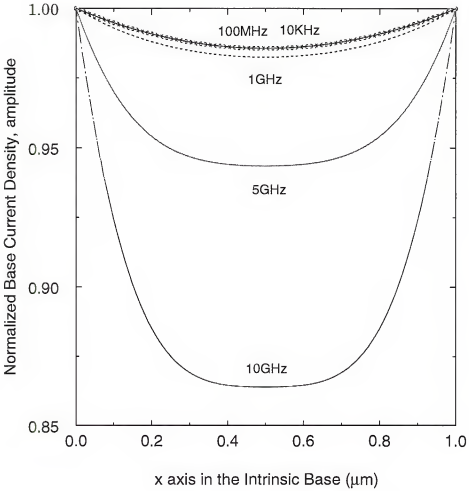


Figure 4.2. The amplitude of normalized base current densities in the intrinsic base region at different frequencies. In all cases, the small-signal voltage biases $v_N = v_F = 3\text{mV}$.

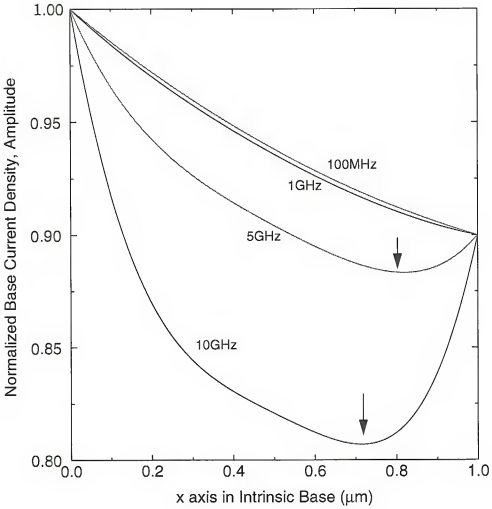


Figure 4.3. The amplitude of normalized base current densities in the intrinsic base region at different frequencies. In all cases, v_N = 3mV and v_F = 2.7mV. The arrows indicate the position of minimum voltage. For the curves of 1GHz and 100MHz, the small-signal voltages decrease monotonically and there is no minimum voltage inside the intrinsic region.

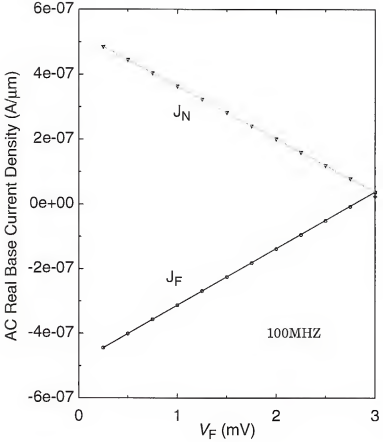


Figure 4.4(a). The real part of the small-signal base current densities plotted as a function of bias v_F when $v_N = 3$mV. The oscillation frequency is 100MHz. Lines are the results of the model and symbols indicate the data from FLOODS simulations.

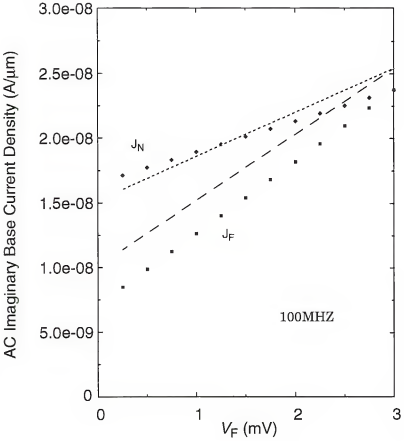


Figure 4.4(b). The imaginary part of the small-signal base current densities plotted as a function of bias v_F when v_N = 3mV. The oscillation frequency is 100MHz. Lines are the results of the model and symbols indicate the data from FLOODS simulations.

Figure 4.5 illustrates the small-signal j-v relation when the device is operated at 10GHz. The imaginary current shows the same behavior as that at the lower frequency operations. Due to the significant decrease of the injection current through the junction, the reduced amount of real J_F is larger than the increase of J_N. Thus, both the total real and imaginary base currents decrease as predicted by the simulations as well as by model. Though there is a numerical discrepancy between the model and simulations, the j-v relation is linear at all frequency ranges of interest, which verifies the validity of eq. (4-13) and (4-14).

4.3 Three Dimensional A.C. Distributed Model

4.3.1 Model Development

The following assumptions are made in the derivation of the distributed model in the base polysilicon region: (i) The capacitance associated with the polysilicon base is neglected since it is in parallel with C_d and is much smaller than C_d. (ii) Since we limit our interests in the biases before high current effects occur, the d.c. voltage bias is approximately constant along FPOLY as can be seen in figure 3.4. (iii) The admittance Y_d is constant everywhere in the intrinsic base region due to the small voltage variation in FPOLY and the intrinsic base region. (iv) The small-signal voltage along lineN in NPOLY is constant due to the small debiasing effect in NPOLY.

The governing equations of the small-signal voltage $v_{PF}(y)$ and current density $j_{PF}(y)$ in FPOLY are given by

$$j_{PF}(y) = j_{PF}(W) + \int_y^W \frac{j_F(y)}{A_P} \cdot dy \tag{4-18}$$

and

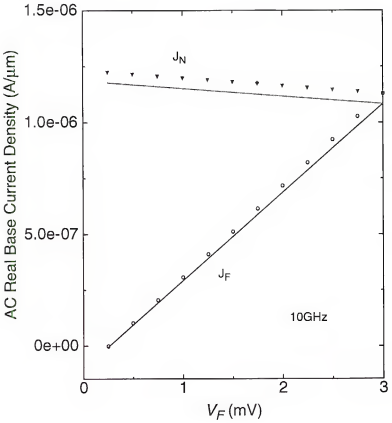


Figure 4.5(a). The real small-signal base current densities plotted as a function of bias v_F when v_N = 3mV. The oscillation frequency is 10GHz. Lines are the results of the model and symbols indicate the data from FLOODS simulations.

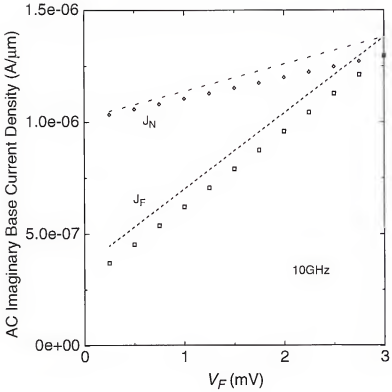


Figure 4.5(b). The imaginary small-signal base current densities plotted as a function of bias v_F when v_N = 3mV. The oscillation frequency is 10GHz. Lines are the results of the model and symbols indicate the data from FLOODS simulations.

$$v_{PF}(y) = v_{PF}(0) - \int_0^y j_{PF}(y)\rho_p \cdot dy \tag{4-19}$$

Following the same methodology used to solve the d.c. problem, the analytical solutions of $v_{PF}(y)$ and $j_{PF}(y)$ can be written as

$$v_{PF}(y) = v_{PF}(0) + \frac{j_{PF}(0)\rho_p}{C_8 \sinh(C_8 W)} \cdot (\cosh((C_8(W-y)) - \cosh(C_8 W))) \tag{4-20}$$

$$j_{PF}(y) = \frac{j_{PF}(0)}{\sinh(C_8 W)} \cdot \sinh(C_8(W-y)) \tag{4-21}$$

where

$$C_8 = \sqrt{\frac{C_5 \rho_p}{A_p}} \tag{4-22}$$

If the current injection from the short edges of emitter stripe is negligible, then the boundary conditions of small-signal voltage $v_{PF}(y)$ and current $j_{PF}(y)$ are given by

$$j_{PF}(0) = \frac{(C_5 - C_6)W}{C_9 A_P} \cdot v_N \tag{4-23}$$

$$v_{PF}(0) = v_N - A_P R_{pe} j_{PF}(0) \tag{4-24}$$

where

$$C_9 = 1 - \frac{C_5 \rho_p}{A_p C_8^2} + C_5 W R_{pe} + \frac{C_5 \rho_p W \coth(C_8 W)}{A_p C_8} \tag{4-25}$$

Note that the only variable in eq. (4-23) and (4-24) is v_N which is the given bias at lineN.

4.3.2 Results and Discussion

The device described in section 3.2.1 with an emitter stripe length of 100μm is used in the analysis. In all cases, d.c. biases Vbe=0.83V and Vcb=2V are applied to the transistor. To improve the accuracy of the terminal j-v relations, the constants C_5 and C_6 in equations (4-13) and (4-14) are obtained directly from the device simulations. Figure 4.6 shows the amplitude of the small-signal voltage along FPOLY at different frequencies while v_N = 3mV. At low frequencies, the base current in the intrinsic base region is dominated by the current conducted through the resistance R_d. This current is small and has no significant debiasing effect in the active and polysilicon base regions. There is also no phase lag due to the insignificant capacitive characteristics of the network as can be seen in Fig. 4.7. The displacement current through C_d increases following the increase of the oscillation frequency becoming more important as it dominates the base current at higher frequencies. The increasing current causes larger ohmic voltage drops along the FPOLY. For instance, the amplitude of the total small-signal current flowing through the FPOLY at 10GHz is 20 times larger than that at 100MHz. The variation of small-signal voltage in FPOLY is more than 72% at a frequency of 10GHz whereas the variation is within 5% when the operation frequency is 100MHz. The phase lag also increases at higher frequencies due to the prominent capacitive characteristics.

The modeled small-signal base current densities injected from the polysilicon contact to the active base region are shown in Fig. 4.8 as a function of position in the base polysilicon. Figure 4.8(a) illustrates the current densities when the operating frequency is 100MHz. At this frequency, the current injection across the junction is approximately constant since the ohmic voltage drop is so small along the base polysilicon as well as in the intrinsic base region. However, the small a.c. debiasing in FPOLY does cause small but asymmetric

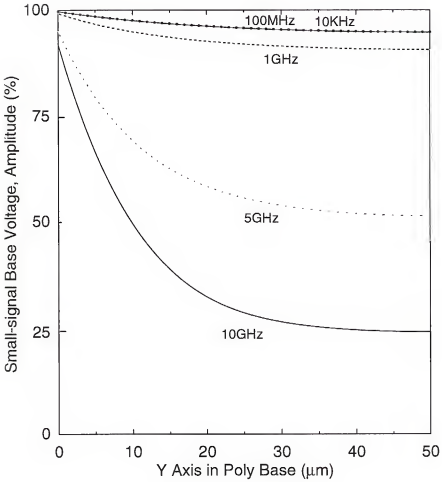


Figure 4.6. The amplitude of the small-signal voltage along FPOLY calculated by the model when v_N = 3mV. The emitter stripe length is 100μm. The voltage and current are assumed to be symmetric to the center of the poly.

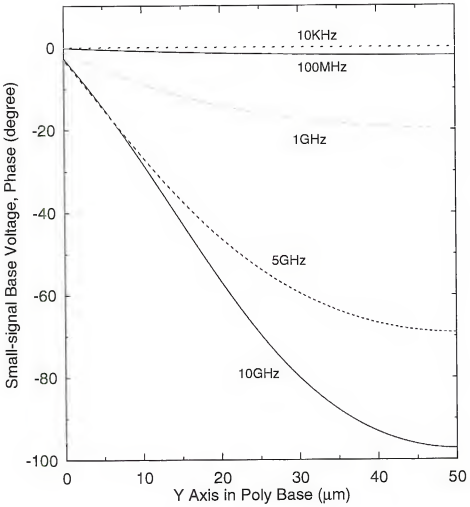


Figure 4.7. The phase lag of the small-signal voltage along FPOLY computed by the model. All the phase lag refer to the voltage v_N.

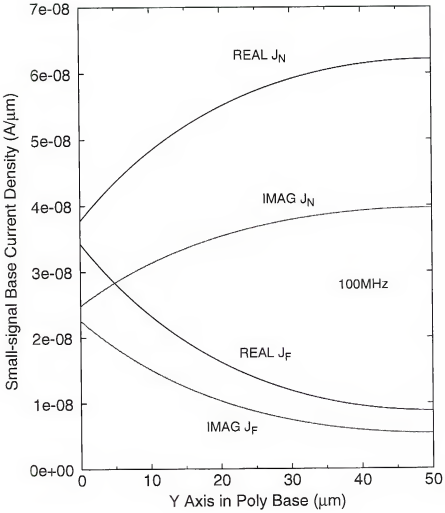


Figure 4.8(a). The small-signal base current densities injected from poly to active base region at a frequency of 100MHz.

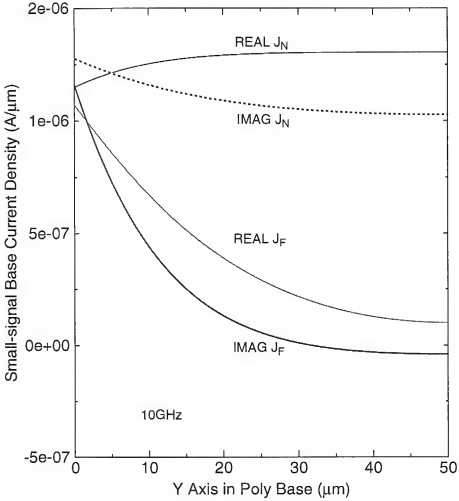


Figure 4.8(b). The small-signal base current densities injected from poly to active base region at a frequency of 10GHz. Both the real and imaginary total current densities decrease as Y increases.

a.c. biases between the base contacts. A similar current dividing phenomenon takes place in the intrinsic base region as that observed in the d.c. asymmetric bias conditions. An obvious nonuniform current distribution is observed in both sides of the base polysilicon contacts and both the real and imaginary currents exhibit the compensation property as seen in the d.c. low current injection cases, *i.e.*, the current lost in the FPOLY is balanced by the increase of current in the NPOLY, and the total current is approximately constant compared with the case of uniform bias in the FPOLY. As the frequency increases, the debiasing effect along the FPOLY becomes significant due to the increasing current resulting in a large decrease of a.c. voltage across the junction in the active base region. Thus, the total current decreases compared with the case of uniform bias in the FPOLY. Figure 4.8(b) shows the current densities at 10GHz. Note the a.c. current at 10GHz is higher than that at 100Mz which results in the considerable debiasing effect in FPOLY. The decrease of real current in the FPOLY is more than the increasing current in the NPOLY and the imaginary currents at both sides decrease toward the center of polysilicon base. So, both the real and imaginary base current calculated by a model that does not take into account the distributed effects in the polysilicon base region will be overestimated. Figure 4.8 also reveals the current redistribution modulated by the operating frequency. The ratio of current contributed by the low resistive NPOLY increases, and that from the high resistive FPOLY decreases as the frequency increases. Thus, in addition to the intrinsic base region, the distributed effect in the base polysilicon region also causes the decrease of complex input impedance as the frequency increases.

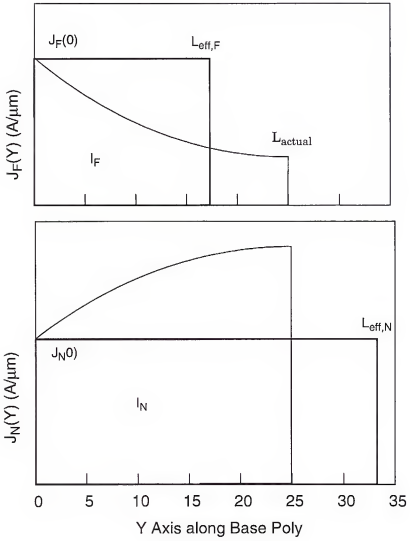


Figure 4.9. The definition of effective length L_{eff} used in this work. The shaded areas indicate the total current $I_{N,F}$ associated with each poly base, which is also equivalent to the product of $L_{eff,N,F}$ and $J_{N,F}(0)$.

4.3.3 Application and Verification

To characterize the high frequency performance of a transistor, 2D device simulations are frequently used. A key assumption of the simulations is that the voltage and current distribution in the third dimension are uniform. The results obtained using our 3D distributed model suggest the assumption may be invalid in certain structures, for some cases, a 3D device simulator is required. However, 3D device simulations are rarely used in device design due to extensive demands on hardware and CPU time. The use of a 2D device simulator coupled with the analytical model to handle the variation of voltage and current in the third dimension is an efficient way to characterize these transistors.

The distributed model is coupled to FLOODS by effective contact lengths. Referring to the base current density plot as shown in Fig. 4.8, the total current $I_{N,F}$ associated with each side of the polysilicon base contact is the integration of the current density $J_{N,F}(y)$ along the polysilicon length (y-direction) which can be equivalent to the current density at the edge of polysilicon contact, $J_{N,F}(0)$, times the effective contact length $L_{eff,N,F}$, *i.e.* $I_{N,F} = J_{N,F}(0) \times L_{effN,F}$. Figure 4.9 illustrates the definition of effective lengths in the NPOLY and FPOLY for the current densities shown in Fig. 4.8(a). L_{eff} of NPOLY is longer than the actual polysilicon contact length L_{actual} since the current density increases along NPOLY. In contrast, L_{eff} of FPOLY is shorter than L_{actual} due to the decreasing current density along FPOLY. Thus, if we obtain the edge current densities $J_{N,F}(0)$ from FLOODS simulations and multiply by the $L_{eff,N,F}$ from the model, we get the actual contact current from the nonuniform 3D effect of the transistor.

Figure 4.10 illustrates the implementation of a realistic 3D structure, as shown in figure 2.14 and 2.15, to a simulated structure. The differences

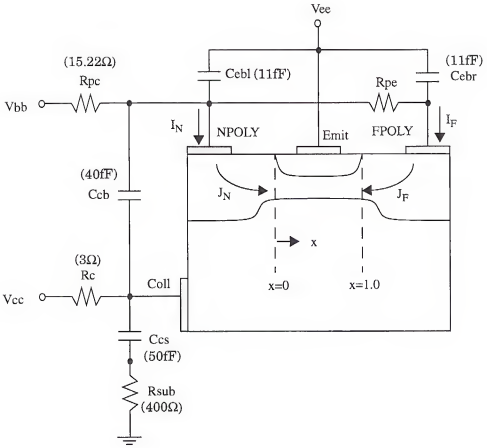


Figure 4.10. Schematic diagram of a BJT with parasitics used in the asymmetric simulations. The total current flowing through each contact of the active device $I_{N,F}$ equals the current density $J_{N,F}$ times the effective length $L_{eff,N,F}$ of the contact. The connection resistance Rpe between NPOLY and FPOLY is 6.3Ω.

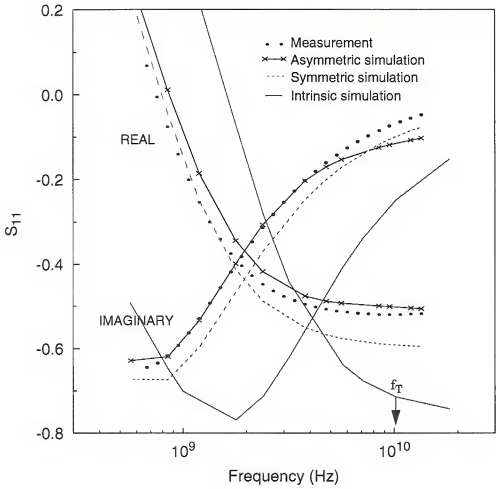


Figure 4.11. The comparison of measured and simulated S_{11} as a function of frequency. The asymmetric simulation uses the structure shown in figure 4.10 and the symmetric simulation utilizes the structure shown in figure 2.16. The cutoff frequency f_T of this device is 10.3 GHz.

between figure 4.10 and 2.16 are that (i) a lumped resistor *Rpe* is placed between the NPOLY and FPOLY contacts to depict the connection resistance between them, (ii) L_{eff}'s are used in the base contacts to describe the nonuniform effect in each contact. Due to the ohmic voltage drop across *Rpe*, the bias at FPOLY is lower than the bias at NPOLY which properly represents the actual asymmetric bias condition at the edge of the poly bases.

To characterize the high frequency performance of a transistor, scattering parameters (s-parameters) are often used because they are easier to measure at high frequencies than other parameters [Sze81]. However, it is difficult to properly describe the transmission line with finite RLC circuit elements in the simulations; therefore, y-parameters are extracted from device/circuit simulations and then converted to equivalent s-parameters [Gon84]. Here we will only focus on the input reflection coefficient S_{11} or input impedance Z_{in} of the device since the base distributed effects have a direct relation with S_{11} and Z_{in} in a common-emitter configuration. S_{11} and Z_{in} are related by

$$S_{11} = \frac{Z_{in} - Z_{o1}}{Z_{in} + Z_{o1}} \tag{4-26}$$

where Z_{in} is the input impedance of the transistor which is associated with the base and emitter characteristics in a common-emitter configuration, and Z_{o1} is the complex characteristic impedance of the transmission line which is given by [Gon84]

$$Z_{01} = \sqrt{\frac{R_0 + jwL_0}{G_o + jwC_0}} \tag{4-27}$$

where R_0, G_0, L_0, and C_0 are the resistance, conductance, inductance and capacitance per unit length of the transmission line. At microwave frequencies, R_0 and G_0 are usually negligible and the transmission line is said to be lossless. In a lossless transmission line, $Z_{o1} = \sqrt{\frac{L_0}{C_0}}$ and is a real. If assuming the induc-

tive component of a transistor is insignificant at the frequency range of interest, then Z_{in} can be represented by a series RC equivalent circuit. At a specific angular frequency w, $Z_{in} = \Re(w) + \frac{1}{jw\chi(w)}$. Substituting into eq. (4-26), it gives

$$S_{11} = \frac{\left((R(w)+Z_{o1})\cdot(R(w)-Z_{o1})+\frac{1}{w^2C(w)^2}\right)-j\cdot\frac{Z_{o1}}{wC(w)}}{(R(w)+Z_{o1})^2+\frac{1}{w^2C(w)^2}} \tag{4-28}$$

The comparison of the measured and simulated S_{11}s are shown in Fig. 4.11 as a function of frequency. The asymmetric simulation uses the structure shown in Fig. 4.10 with the effective contact lengths in the device/circuit interface which take into account the nonuniform effects in the base polysilicon region. The symmetric simulation utilizes the same structure with the actual contact lengths in the interface and R_{pe}=0 which results in a symmetric bias at the NPOLY and FPOLY. For the intrinsic simulation, the circuit network is not included. It is obvious that the simulation cannot predict correctly if the parasitics are not being accounted for. The real part of S_{11} is reasonably described by both the asymmetric and symmetric simulations at the low frequencies. As the frequency increases, the result of the symmetric simulation deviates from the measured data, whereas the asymmetric simulation result remains in agreement with the measured one. For imaginary S_{11}, there is a discrepancy between measurement and symmetric simulation, while the asymmetric simulation result agrees with the measured data over the low frequency range, and as the frequency increases the measured results deviate from the asymmetric simulations.

An analysis of the input impedance Z_{in} reveals that $\Re(w)$ is a weak function of frequency, but $\chi(w)$ is a strong function of frequency. From Eq. (4-28), it is clear to see that the $w\chi(w)$ term dominates the real part of S_{11} at low

frequencies and the $\Re(w)$ term becomes as important as $w\chi(w)$ at intermediate frequencies. At high frequencies, the real S_{11} is dominated by the $\Re(w)$ term. The imaginary part of S_{11} is governed by the $w\chi(w)$ term over all frequency ranges. From these results we can conclude that the resistive characteristics of the uniform transmission line model in the active base region is properly modeled if the distributed effect in the polysilicon base region is included. The model can also accurately describe the capacitive property for low and medium frequency applications and becomes less accurate at high frequencies. This discrepancy may be due to the nonuniform capacitance in the intrinsic base region which in fact is not rectangular resulting in a nonuniform doping profile in the base-emitter junction along the x-direction.

$|S_{11}|^2$ represents the ratio of the power reflected from input port to the power available at the input port and is associated with the maximum available power gain and stability factor of a transistor. At the cutoff frequency f_T (= 10.3 GHz), the $|S_{11}|^2$ predicted by the asymmetric and symmetric simulations is higher than the measured data by 3% and 32%, respectively.

4.4 Summary

An a.c. model has been developed to study the high frequency distributed effect of voltage and current in the intrinsic and polysilicon base regions. The small-signal voltage and current distributions in both regions have been described completely by the model. By the introduction of the effective contact length, the model has been applied to a 2D device simulator for quasi-3D simulations. S_{11} was extracted from different simulations and compared with the measured data. The agreement between measured and simulated S_{11} has been improved by the consideration of the nonuniform effect in the third dimension of the device. A modification of the nonuniform capacitive character-

istic in the small-signal network is required to further improve the accuracy of the imaginary S_{11} at high frequencies.

CHAPTER 5
VERIFICATION OF DC DISTRIBUTED MODEL AND GEOMETRY DEPENDENT DEVICE PERFORMANCE DUE TO DISTRIBUTED EFFECTS

5.1 Introduction

Stripe and interdigitated configurations (Fig. 5.1) are two popular layout geometries used in high speed BJT design. Designers usually assume the current and voltage distribution in the z-direction of the polysilicon contacts are uniform, and the large- and small-signal currents are linearly proportional to the length of the stripes or fingers. The analysis of Ch. 3 and Ch.4 shows this assumption is not correct for striped single metal base contact structure (SMB) in some applications. Similar debiasing and distributed effects also take place in the fingers of interdigitated structures. However, DC behaviors of varying types of layouts and sizes can differ due to varying degrees of influence by distributed effects on the different structures.

The influence of the distributed effects on the structures strongly depends on the sheet resistance of the base polysilicon contacts. A current path of low resistance reduces debiasing and diminishes distributed effects. Silicided polysilicon technologies are often utilized in current high speed semiconductor devices to reduce the resistivity and RC delay. The sheet resistance of contemporary silicide technologies ranges from single to teens Ω/sq [Wol86, Chi91, Yam93]. A low resistance silicide process usually increases the complexity of fabrication and cost. Thus, the adequate choice of the sheet resistance of polysilicon contacts is one of the key issues in the device design.

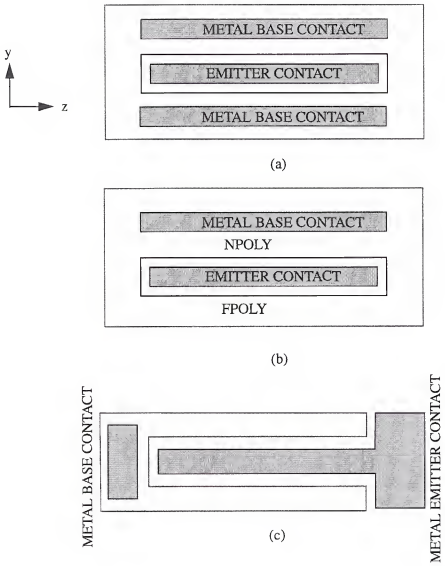


Figure 5.1 Three layout structures for high speed BJTs. (a) Double Metal Base contact structure (DMB). (b) Single Metal Base contact structure (SMB). (c) Interdigitated structure.

Three-dimensional simulation capability has been implemented into FLOODS allowing a direct study of the 3D distributed effects. In this chapter, the 3D simulations are demonstrated and some questions about 3D simulations are discussed. The results of the DC distributed model derived in Ch. 3 are compared with results from 3D simulation to verify the accuracy and limitations of the model. The device performance of double polysilicon BJTs with DMB, SMB, and interdigitated layout structures is studied for different geometry sizes and polysilicon contact sheet resistances.

5.2 3D Device Simulations and Simulation Structures

The box's (finite difference) method is used as the discretization scheme for the 3D simulations. The algorithm to solve a 3D problem is the analogous as that of 2D simulations. Using the discretized form of Poisson equation as an example,

$$\sum_j \vec{D}_{ij} \cdot dS_{ij} = \rho_i \cdot dV_i \tag{5-1}$$

where $\vec{D}_{ij}$ is the potential flux along edge ij between nodes i and j, and ρ_i is the net charge associated with node i. In 3D simulations, dS_{ij} is the cross section area bisecting edge ij, and dV_i is the volume associated with node i, while dS_{ij} is the cross section length bisecting edge ij, and dV_i is the area related to node i in 2D simulations. For object-oriented languages like C++, the program can automatically choose the proper dimensions for dS_{ij} and dV_i in 2D or 3D simulations. Therefore, only a small amount of additional coding is necessary to accommodate the 3D capability into FLOODS.

One main difference between the 2D and 3D simulations is the density of the sparse matrix. For a node in a 1D mesh, there are two adjacent nodes.

These adjacent nodes become six in a 2D right triangular mesh and increase to 14 in a 3D tetrahedral mesh. Each adjacent node represents a nonzero entry in one row of the matrix. Therefore, the sparse matrix of 3D simulations is denser than the matrix of 2D simulation for the same number of nodes. For example, the number of nonzero entries of a drift-diffusion 2D simulation with 2160 nodes is 99218. For a similar 3D simulation with 2080 nodes, the number of nonzero entries becomes 163294 which requires twice as much memory and CPU time as compared with 2D simulations.

The 3D simulation structure of a high speed npn BJT is illustrated in Fig. 5.2. The material of the bottom cube is silicon and the two sections stacked on the silicon are p-type base polysilicon contacts. For a DMB layout structure, the metal contacts are located at faces A and B and are connected to a voltage source. The metal contacts of a SMB structure are made on face A and D where the left section of polysilicon acts as NPOLY and the right section of polysilicon is FPOLY. There is a resistor connected between the two contacts to represent the resistance of polysilicon, R_{pe}, between NPOLY and FPOLY, and Contact A is connected to a voltage source. The metal contacts of the interdigitated structure are made on faces C and D and are connected to a voltage source. The sheet resistance of the polysilicon contact R_{sh} is controlled by the specified constant hole mobility in the polysilicon material which is independent of the mobility in the silicon region.

The accuracy and convergence capabilities of device simulations strongly depend on the mesh density and location associated with the device structure. The regions where large variation of doping, potential and carrier densities, etc., exist require fine grid to properly describe these physical parameters at the specific positions of the regions. Figure 5.3 illustrates the Gummel plot of a 5μm stripe length npn BJT for varying mesh densities in the 3D

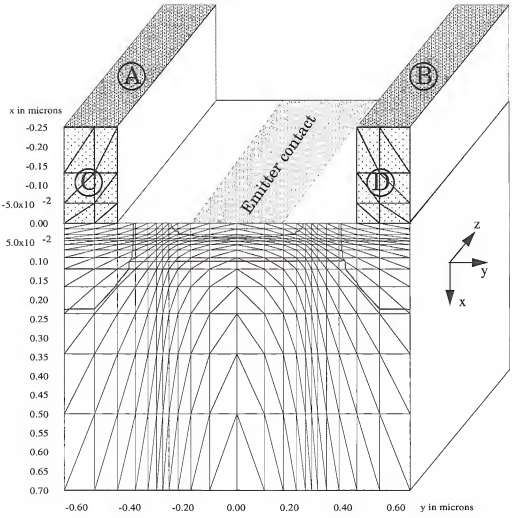


Figure 5.2 The 3D simulation structure of a BJT. A, B, C and D mark the faces where a metal contact is attached to the polysilicon.

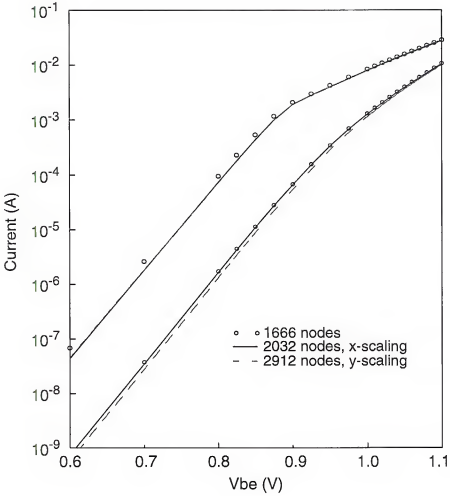


Figure 5.3 The Gummel plots of a BJT simulated with a varying number of grid points in the mesh structure.

simulations. The first mesh consists of about 400 nodes in each xy-plane with 4 xy-planes (3 segments) distributed at 5μm intervals along the z-direction. There are 1666 nodes in this mesh. The second mesh refines the grids in the x-direction and the number of nodes becomes 2032. Due to the finer grid in the base-collector depletion region, the second mesh predicts a wider base width. Thus, the collector current of the finer mesh structure is 9% smaller at V_{be}=0.7V compared with the 1666 node mesh. The base currents of these two meshes do not show significant differences since the doping density and depletion region boundaries around the base-emitter junction is not significantly changed by the finer mesh. The third mesh further scales the grid in the y-direction. In this case, the base doping at the sidewall of the base-emitter junction decreases, reducing the injection of holes into the emitter. Therefore, the base current is decreased by 25% and 20% at V_{be}=0.7V compared with the first and second meshes, respectively. Since few electrons are injected from the highly doped sidewall into the base, the collector current remains unchanged.

Further scaling of the mesh in the xy-planes improves the accuracy, but this improvement becomes negligible after the mesh is fine enough. If the doping profile in the z-direction is uniform, the accuracy of the device simulations is insensitive to the mesh density. For example, the use of three segments or five segments of mesh does not show significant difference for the I-V curve of a SMB device where a significant debiasing occurs in the z-direction. This is probably due to the smooth change of voltage which can be properly approximated by these segments. However, it is observed that the use of a very large grid size in the z-direction reduces the convergence capability due to the large aspect ratio in the tetrahedral unit of the mesh. Therefore, if the piece-wise linear voltage can properly approximate the real voltage in the z-direction and

convergence is not a problem, the grid size in the z-direction can be set as large as possible.

Limited by the available computer resources (Sun Ultrasparc 2 workstations with 1.6G accessible memory), all the 3D simulations in this and the following chapters use the first mesh (400 nodes in xy-plane as shown in Fig. 5.2). The grid size in the z-direction varies from one to 2.5μm depending on the length of the device. The total number of nodes in the meshes is limited to 2100 nodes which requires about 1.5G of memory for AC simulations. Though the loose mesh in the xy-plane introduces error in the I-V curves, the trend and the characteristics of the distributed effects in the transistors can still be properly described by this 3D simulations.

Usually a base polysilicon contact covers over both the base diffusion region and the field oxide region. The sheet resistance on the top of field oxide is typically higher than that on the diffusion region. To simplify the simulated structure, the nonuniform sheet resistance in the y-direction can be represented by a single effective sheet resistance R_{sh} which is obtained by simulating a complete 3D polysilicon structure including the base diffusion and field oxide region: the resistance of polysilicon R is computed from the terminal voltage and current, then $R_{sh} = R \cdot \frac{W_{poly}}{L_{poly}}$, where W_{poly} and L_{poly} are the width and length of the polysilicon, respectively. This effective sheet resistance can properly describe the debiasing effects, although the simulation structure as shown in Fig. 5.2 does not include the entire polysilicon structure.

5.3 DC Characteristics

As discussed in section 3.3.3, FPOLY in the SMB structure represents a highly resistive current path and thus has a bigger effective base resistance than the DMB structure. For the same reason, both fingers of the interdigitated

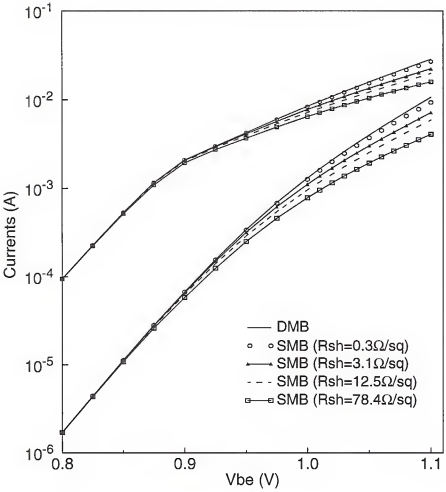


Figure 5.4. (a) The Gummel plots of BJTs with DMB and SMB layout structures obtained by 3D simulations. Four different R_{sh}s of base polysilicon are used in the simulations for the SMB structure.

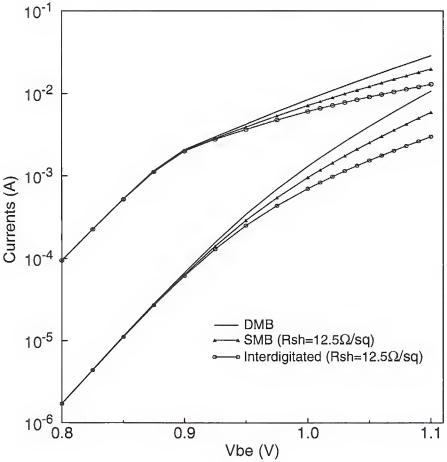


Figure 5.4. (b) The simulated Gummel plots of BJTs with different layout structures. The value of R_{sh} in polysilicon contacts is 12.5Ω/sq.

structure have highly resistive current paths compared with the paths in the DMB device. Therefore, the DMB structure has the lowest effective base resistance, and the interdigitated structure has the highest one. The effective base resistance can be calculated from the deviation of base current from the ideal exponential curve in a Gummel plot. The difference of the voltage is mainly due to the ohmic voltage drop in the base and emitter regions as expressed in Eq. (2-18). Qualitatively, a larger degree of base-current-curve-bending in a Gummel plot indicates a larger base resistance.

The Gummel plots of a 5μm striped BJT with different layout structures and R_{sh}s are shown in Fig. 5.4. At low biases (< 0.85V), the debiasing is very small either in the active base or in polysilicon base contacts. The base current increases exponentially with V_{be} and is independent of layout structures and R_{sh}. As the bias increases, the base current curves for different layouts and R_{sh} depart from each other. It is clear that the base resistance of the SMB structure is always larger than that of the DMB, and a larger sheet resistance increases the difference. For the same value of R_{sh}, the degree of bending for the interdigitated structure is larger than that of the SMB transistor as plotted in Fig. 5.4(b). Therefore, the layout with the DMB structure achieves the lowest base resistance, and the interdigitated structure gives the highest base resistance.

Figure 5.5 shows the hole Fermi potential along FPOLY of a SMB device at V_{be}=0.95V computed by 3D simulations and models. The current density in the xy-plane of FPOLY in the model is assumed to be constant. However, the current density in FPOLY of the 3D simulations is not uniform due to the nonuniform doping density in xy-plane. Thus, for the purpose of comparison, two potential curves are extracted from the top surface of FPOLY, one from the inner side (y=0.45μm) and the other from the outer side (y=0.65μm). As seen in Fig. 5.5, the potential in FPOLY descends monotonically from the contact side

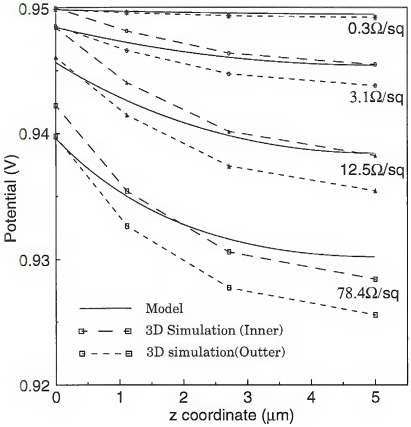


Figure 5.5 The potential in FPOLY of a SMB structure at a bias of 0.95V. The metal contact is located at z=0 which is attached to a voltage source by a resistor *Rpe*. The inner potential is extracted from the inner side (y=0.45μm) of the polysilicon top surface in the 3D simulation structure. The outer potential is from the outer side of the polysilicon top at y=0.65μm.

(z=0) toward the center (z=5μm), and a higher R_{sh} causes a larger debiasing. There is also a connecting resistance R_{pe} between base polysilicon contacts resulting in additional ohmic voltage drop at the edge of FPOLY. For instance, the potential difference between the edge and center of a FPOLY with R_{sh}=12.5 Ω/sq is about 10mV and voltage drop across R_{pe} is about 4mV obtained from the simulations. This debiasing results in a significant decrease of the base current compared with the DMB structure as seen in Fig. 5.4. The curves predicted by the model follow the same trend and basically fall within the range of the simulations. Though there is a discrepancy, the current from each contact computed by the model is about the same as that obtained by the 3D simulations.

From the analysis in section 3.2.4, the biases in two base polysilicon contacts of the DMB structure are constant along z-direction and equal to each other. The debiasing simply takes place in y-direction symmetrically and the currents flowing into each base polysilicon contact are the same. For a SMB structure, the voltage in NPOLY is uniform but voltage decreases along FPOLY resulting in asymmetric biases in the two polysilicon base contacts. Consequently, the currents from two contacts become asymmetric as well. Figure 5.6 illustrates the simulated potential profile in the intrinsic base region beneath the emitter of a SMB structure with R_{sh}=12.5 Ω/sq. It is observed that debiasing in the intrinsic base region results in a voltage decreasing toward the center in the y-direction. The voltage decrease in FPOLY also causes the potential to decrease in the z-direction. Consequently, the minimum voltage point moves to the right (positive y) and toward the center of the stripe (Z=5μm) due to the potential decrease in FPOLY. Figure 5.6 shows exactly the same characteristics as Fig.3.5 computed by the model.

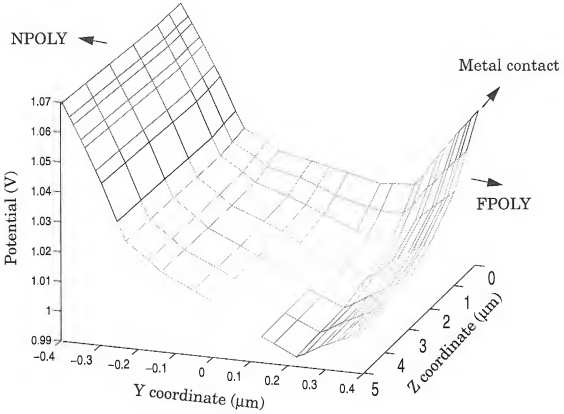


Figure 5.6 The potential in the intrinsic base region of a SMB transistor obtained from the 3D simulations at a bias of V_{be}=1.1V.

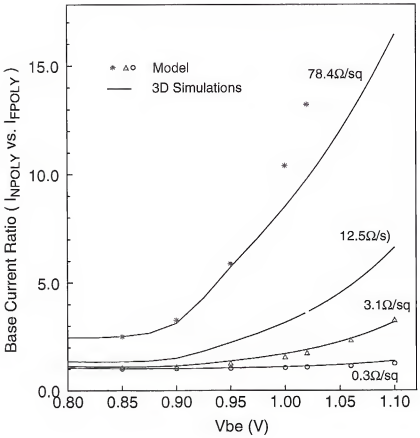


Figure 5.7 A comparison between simulated and modeled base current ratio as a function of V_{be} at four different values of R_{sh}.

The asymmetric biases at the base polysilicon contacts cause the movement of the minimum voltage point in the intrinsic base region resulting in unequal current contributions from NPOLY and FPOLY. The ratio γ of the current contributed by NPOLY vs. current from FPOLY of SMB structures is greater than unity since the voltage at NPOLY is always higher than the voltage of FPOLY. Fig. 5.7 illustrates the simulated and modeled γ as a function of bias and R_{sh}. For a very low R_{sh} device (< 1Ω), the ratio γ approaches unity since the voltage difference between these two polysilicon contacts is negligible even at high biases. The debiasing is also not significant in the devices with intermediate R_{sh} ($1\Omega < R_{sh} < 15\Omega$) at a bias less than 0.9V. The further increase of V_{be} causes larger ohmic voltage drops along FPOLY and enlarges the voltage difference between NPOLY and FPOLY which enhances the movement of the minimum voltage point toward FPOLY side in the intrinsic base region. Consequently, the current from NPOLY increases, and the current from FPOLY decreases, while the ratio γ becomes much greater than one. The increase of R_{sh} tends to trigger the movement of the minimum voltage point at a lower bias and induces larger movement at the same bias compared with the lower R_{sh} devices. Thus, a very unbalanced current contribution from the two polysilicon contacts is observed. Since the current contribution from FPOLY is negligible at high bias for larger values of R_{sh}, the performance of the SMB device is similar to that of a single base polysilicon contact device.

Figure 5.7 shows that the model predicts the current ratio γ accurately for devices with low and intermediate R_{sh}. The modeled γ is also well matched to the simulation results of devices with higher R_{sh} until it goes into the high current region. In this region, the model overestimates the ratio γ. The error comes from the failure of the linearized assumption in Eq. (3-10) where the voltage difference between two base polysilicon contacts is assumed to be much

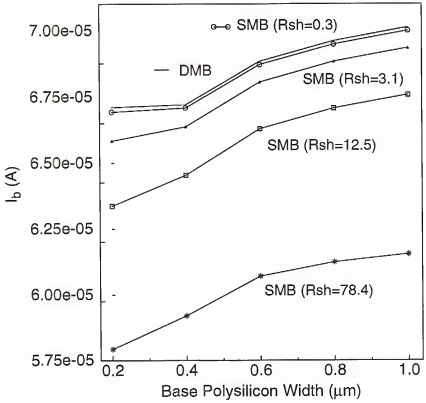


Figure 5.8 The base current plotted as a function of base polysilicon width and sheet resistance obtained from 3D simulations.

smaller than two times the thermal voltage V_o (~52mV at 300K). For those devices with large R_{sh} operating in the high current regime, the voltage difference between the base polysilicon contacts approaches $2V_o$. Under this condition, the increase of J_{NPOLY} following the decreasing V_F is less than that predicted by the linearized model due to the additional voltage drop in the base region. Consequently, there is less I_{NPOLY} in the real situation resulting in the overestimation of γ by the model.

Figure 5.8 shows the base currents of SMB and DMB structures with different polysilicon contact widths W_{poly}. The wider contact reduces the resistance and decreases the ohmic voltage drop along the base current path. Thus, the base current (and also the collector current) of devices with wider W_{poly} is slightly larger than that of a narrower W_{poly}. However, the use of wider W_{poly} does not significantly eliminate the distributed effects. R_{sh} is still the dominant parameter in determining the debiasing and distributed effects. Though a narrow W_{poly} is used for the devices studied in Ch. 5 and 6, the same results are applicable to wider W_{poly} devices.

A circuit designer usually assumes the current driving capability (current per unit length) is linearly proportional to the length of the device stripe. If twice the amount of current driving is required, the layout stripe length is simply doubled. Figure 5.9 plots the base current driving capability as a function of stripe length for various layout structures and sheet resistances. The base current of the DMB structure is seen to be linearly proportional to the stripe length. For the SMB and interdigitated structures, the unity-length current driving capability degrades as the stripe length increases, and the larger R_{sh} causes more significant decay due to a larger voltage decrease in the base polysilicon. For the same R_{sh}, the degradation of the interdigitated structure is larger than the SMB structure since both base

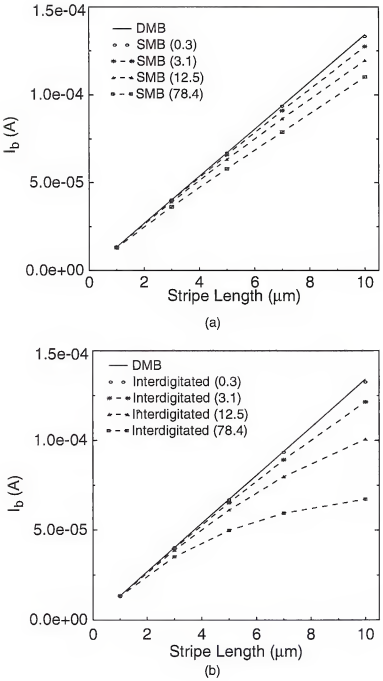


Figure 5.9 The base current vs. stripe length. (a) Comparison between DMB and SMB structures. (b) Comparison between DMB and interdigitated structures.

polysilicon contacts of the interdigitated structure suffers from the debiasing effects. The collector current driving capability also shows the same characteristics as that of base current illustrated in Fig. 5.9. Since the input and output current driving is very sensitive to R_{sh}, the silicided process is necessary and very important to the DC characteristics of the DMB and interdigitated structures.

5.4 Summary

The base resistance of BJTs is shown to strongly depend on the 3D distributed effects. Due to different degrees of influence by the effects, DMB, SMB and interdigitated layout structures present different I-V relations and have different values of base resistance. The current driving capability is also affected by the distributed effects. The DMB layout is shown to have the best current driving capability and lowest base resistance, while the interdigitated structure has the largest base resistance and the worst driving capability for the same process.

Three-dimensional device simulations have been used to verify the accuracy and limitation of the quasi-3D distributed model. The results show the model is valid if the voltage difference between the two polysilicon base contacts is smaller than two times the thermal voltage. The condition is violated only if the device has a large polysilicon sheet resistance and operates at high biases where there is less interest in practical device applications.

CHAPTER 6
VERIFICATION OF AC DISTRIBUTED MODEL AND GEOMETRY DEPENDENT AC DEVICE PERFORMANCE DUE TO DISTRIBUTED EFFECTS

6.1 Introduction

The small-signal current flowing in the intrinsic and polysilicon base regions of a BJT also represents the distributed characteristics and build-up of small-signal debiasing along the current paths. The input conductance of common-emitter BJTs increases following the increase of operating frequency due to the larger input small-signal current flowing through the capacitance associated with base. Therefore, more significant debiasing effects are observed at higher frequencies. For a SMB structure, the increasing small-signal current creates larger asymmetric biasing resulting in unbalanced small-signal current drive in the two polysilicon base contacts. For an interdigitated structure, the biases in the base contacts are symmetric and the debiasing takes place in both base fingers. Thus, the use of various layout geometries can result in different high frequency characteristics for the same device process.

In this chapter, the AC distributed model derived in Ch. 4 and the results of quasi-3D simulations are verified by comparing them with the 3D simulation results. The use of various sheet resistances R_{sh} in the base polysilicon contacts on the AC performance of BJTs is investigated. The AC characteristics of different layout structures and geometry sizes are also studied based on the 3D simulations and the importance of the distributed effects on the RF circuit design is demonstrated as well.

6.2 Verification of AC Model and Quasi-3D Simulations

The AC performance of BJTs with various layout structures are studied in this chapter and characterized at biases of Vbe=0.85V and Vcb=2V where the base-push out effect has not taken place, and DC debiasing does not significantly affect the base and collector currents (refer to Fig. 5.4). Figure 6.1 shows the modeled small-signal voltage in FPOLY of a SMB structure characterized at four frequencies when a small-signal voltage is applied to the base voltage source. It is observed that the magnitude of the small-signal voltage decays along FPOLY and the debiasing is larger at higher frequencies due to the increasing small-signal current level. There is also a phase lag in the voltage along FPOLY and the degrees of delay following the increase in frequency. At very high frequencies, the phase delay near the front side of FPOLY is less than that of lower frequencies. The reason for the smaller delay is the current crowding at the edges of the intrinsic base region near the polysilicon contacts which reduces the effective current paths and decreases the RC delay near the metal contact. However, the delay away from the metal contact is still larger than that at lower frequencies due to the delay in FPOLY itself.

Though small, the slight decay of the voltage in FPOLY at low frequencies causes asymmetric biases and moves the minimum voltage point l toward FPOLY. The point l acts as a current divider in the intrinsic base region which forces part of the current originally flowing through FPOLY at symmetric bias to flow through NPOLY. Consequently, the current I_{NPOLY} from NPOLY increases with the decrease of current I_{FPOLY} through FPOLY. The contribution of current from each base polysilicon contact can be characterized by the effective length L_{eff} defined in section 4.3.3. If L_{eff} is longer than the actual polysilicon contact length L_{actual}, the current contributed from that contact is

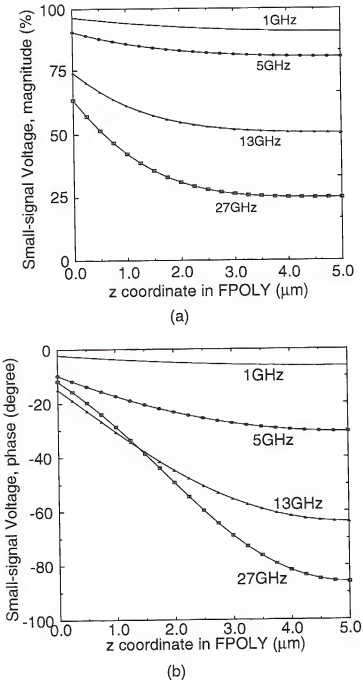


Figure 6.1 The modeled small-signal voltage in FPOLY of a 5μm SMB transistor when a small-signal voltage with a magnitude of 3mV is applied to the base voltage source. (a)Magnitude. (b)Phase.

more than the current contributed from the same contact while the biases are symmetric, and vice versa.

The comparison of L_{eff} from the model and 3D simulations are plotted in Fig. 6.2 as a function of frequency and R_{sh}. At low frequencies, L_{eff} of NPOLY is longer and L_{eff} of FPOLY is shorter than L_{actual} (=5μm). The use of higher R_{sh} increases the difference. However, the sum of the two L_{eff}s approximately equals two times L_{actual} due to the division of the constant base current between the two polysilicon contacts when the debiasing in the intrinsic base region is negligible. At intermediate frequencies, the magnitude of the small-signal voltage in FPOLY decreases slightly while the phase delay results in a shift from imaginary to real current. Thus, the real L_{eff}s become longer and the imaginary L_{eff}s are shorter for both sides of the polysilicon contacts. As the operating frequency further increases, both the real and imaginary L_{eff}s decrease for both contacts due to the large debiasing in FPOLY. The use of smaller R_{sh} in polysilicon contacts reduces the influences of debiasing effects at all frequencies. Therefore, the L_{eff}s of both NPOLY and FPOLY are close to L_{actual} compared with devices having larger R_{sh}. Good matches are obtained between the modeled and simulated L_{eff}s at all frequencies for different R_{sh}s verifying the accuracy and validity of the 3D distributed model.

Quasi-3D simulations are performed using the coupling of 2D simulations and L_{eff}s obtained in Fig. 6.2, and the resulting Y_{11}s are compared with that of 3D and 2D simulations as shown in Fig. 6.3. For the DMB structure, Y_{11} from the 2D simulation is the same as Y_{11} of the 3D simulation since the voltage distribution in the polysilicon contacts of the device are uniform. Y_{11} of the SMB structure is similar to that of the DMB at low frequencies. The use of 2D simulation can accurately predict the input characteristics of the SMB devices in this frequency range. However, Y_{11} of the SMB structures deviates

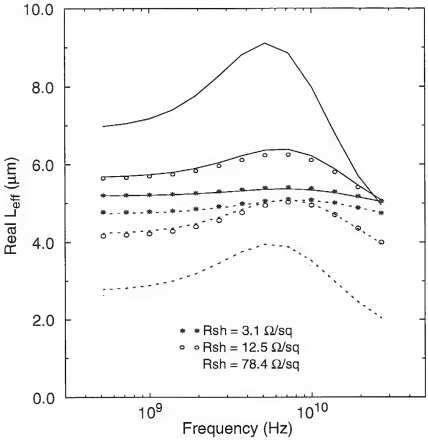


Figure 6.2. (a) The real effective lengths of NPOLY and FPOLY in a SMB device at different frequencies. The symbols are obtained from 3D simulations and lines are computed by quasi-3D distributed model. Solid lines indicate the L_{eff} curves of NPOLY and dashed lines are L_{eff} curves of FPOLY.

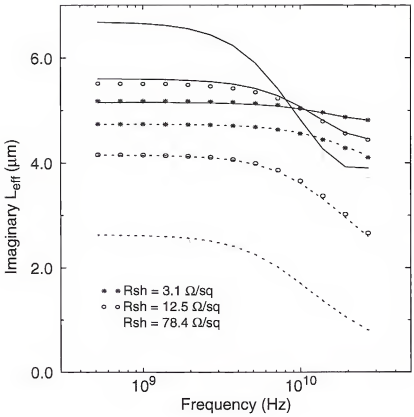


Figure 6.2. (b) The imaginary effective lengths of NPOLY and FPOLY in a SMB device at different frequencies. The symbols are obtained from 3D simulations and lines are computed by model. Solid lines indicate the L_{eff} curves of NPOLY and dashed lines are L_{eff} curves of FPOLY

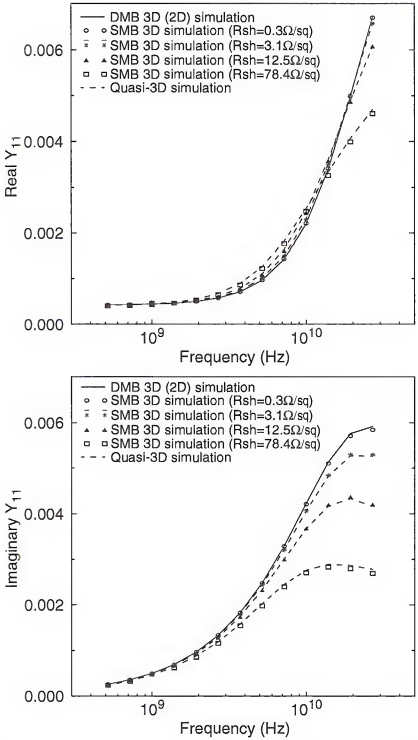


Figure 6.3 A comparison of simulated and modeled Y_{11} with DMB and SMB layout structures.

from that of the DMB at high frequencies due to the debiasing and phase delay in FPOLY. The higher the sheet resistance, the larger the debiasing and phase delay, resulting in a bigger discrepancy between Y_{11} of the SMB and DMB structures. 2D simulation can no longer properly describe the input behavior of a SMB transistor at this condition. In contrast, quasi-3D simulations match the results of 3D simulations at all frequency ranges of interest for various values of R_{sh}. Therefore, the use of quasi-3D simulations improves the accuracy of 2D simulations and is more efficient than 3D simulations.

6.3 Structure and Geometry Dependent AC Performance

The distributed effects are also important to other y-parameters and show different degrees of influences on AC behaviors with different R_{sh}, layout structures, and geometry sizes. Figure 6.4 illustrates the deviation of Y_{21} of the SMB devices relative to Y_{21} of DMB structure. The use of higher R_{sh} decreases real Y_{21} in SMB due to larger debiasing in FPOLY. Higher R_{sh} also causes larger small-signal voltage phase delay in FPOLY and increases imaginary Y_{21} at intermediate frequencies. The imaginary Y_{21} of higher R_{sh} devices decreases faster at higher frequencies due to the larger voltage drop in FPOLY following the increase in frequency.

The backward transmission current mainly passes through the base-collector junction capacitance C_{jbc} and then divides into two components, one flows through the emitter and another through the base. Varying R_{sh}s in SMB devices does not affect Y_{12} at low frequencies since only a small amount of current passes through C_{jbc}. As the frequency increases, more current flows through C_{jbc}. The device with higher R_{sh} has a larger effective base resistance which reduces the current passing through it. Consequently, Y_{12} of those devices becomes smaller at high frequencies as shown in Fig. 6.5. Larger effec-

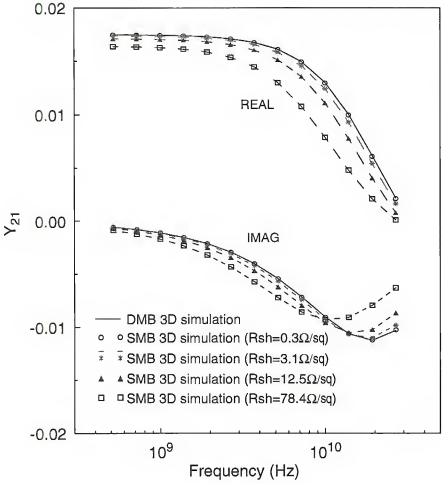


Figure 6.4 The Y_{21} plot of DMB and SMB structures with different R_{sh} obtained from 3D simulations.

tive base resistance also increases the potential in the base region forcing more electrons to be injected from the emitter into the collector. Thus, Y_{22}s of the devices with larger R_{sh} are larger than that of smaller ones at low and intermediate frequencies as illustrated in Fig. 6.6. The collector current becomes dominant by the current flowing through C_{jbc} as the frequency increases. The larger effective base resistance in SMB structures induce less phase angle which increases the real and decreases the imaginary collector currents, and likewise the real and imaginary part of Y_{22}. Though the use of different values of R_{sh} significantly changes the y-parameters, the cut-off frequencies f_T of these devices (~32GHz) are the same since f_T is mainly determined by the properties of the base-emitter junction and the intrinsic base regions.

The distributed effects acting in FPOLY of the SMB structure strongly influence the small-signal current behavior at high frequencies. For the interdigitated structure, similar distributed effects take place in both base polysilicon contacts. The modulation of current at high frequencies is more significant than in the SMB devices since debiasing occurs in both fingers of interdigitated structure. Figure 6.7 compares Y_{11} for the DMB, SMB and interdigitated structures with R_{sh}=12.5Ω/sq. It is seen that both the real and imaginary parts of Y_{11} of the interdigitated structure are smaller than that in the SMB device at high frequencies, and the deviation from Y_{11} of DMB is more significant compared with SMB devices. Thus, the characterization of the interdigitated structure with 2D simulations results in a more significant error than the SMB structure.

The input and output small-signal current driving capability is usually assumed to be linearly proportional to the stripe length, and the phase delay in the polysilicon contacts are neglected. For the DMB structure, there are no distributed effects in the polysilicon contacts. Therefore, $|Y_{11}|$ and $|Y_{21}|$ are

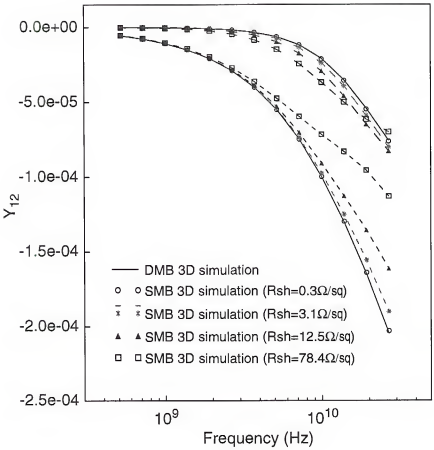


Figure 6.5 The Y_{12} plot of DMB and SMB structures with different R_{sh} obtained from 3D simulations.

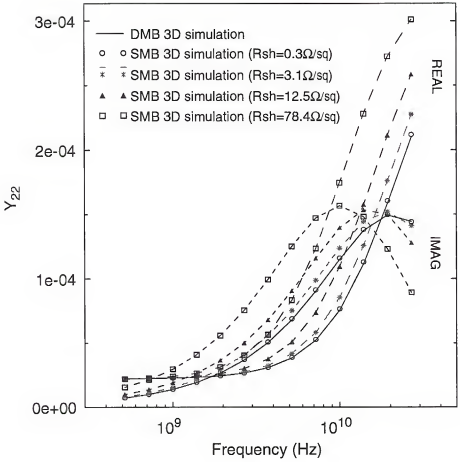


Figure 6.6 The Y_{22} plot of DMB and SMB structures with different R_{sh} obtained from 3D simulations.

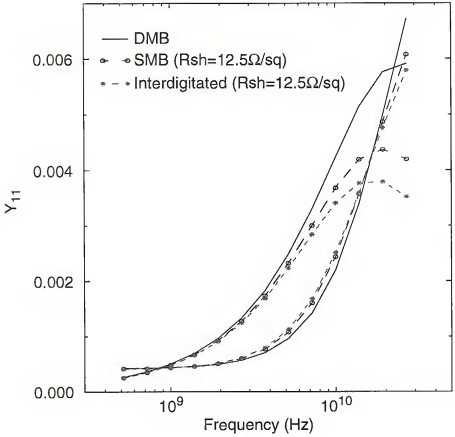


Figure 6.7 A comparison of Y_{11} between DMB, SMB and interdigitated structures with R_{sh}=12.5Ω/sq obtained from 3D simulations.

linearly proportional to the stripe length as shown in Fig. 6.8.(a) and (b), respectively. The phase delays are also the same regardless of the length of the stripe used as illustrated in Fig. 6.9. However, in the presence of distributed effects, debiasing along FPOLY in the SMB and the fingers in interdigitated structures reduces the input and output current drive and degrades the linearity at high frequencies. The signal propagating along FPOLY and the fingers also requires a longer time compared with that of the DMB structure resulting in additional phase delay. A longer stripe further degrades the current drive capability and enlarges the RC delay unless a very low value of R_{sh} is utilized. The sheet resistance of the base polysilicon contact plays a key role in determining current driving capability and phase delay in SMB and interdigitated devices. For example, the input current driving capability of a 10μm SMB device with R_{sh}=12.5Ω/sq is 18% less and phase delay is 12° more than a DMB structure. If a R_{sh} of 3.1Ω/sq is used, the current loss and additional RC delay reduces to 9.3% and 5°, respectively. Thus, a good silicidation in the base polysilicon contacts can significantly improve the characteristics of SMB and interdigitated BJTs.

The interdigitated structure has worse current driving capability and larger phase delay than the SMB device since the debiasing effects and RC delay take place in both polysilicon contacts. Due to distributed effects in the base polysilicon, at its worst, the current in a SMB device is totally supplied by NPOLY. For the interdigitated structure, the distributed effects keep affecting the current and phase delay as the frequencies and stripe length increase. Consequently, the input and output current drive of SMB devices increase monotonically, while the current supply of interdigitated devices tends to saturate as the stripe length increases. At the same time, the phase delay of

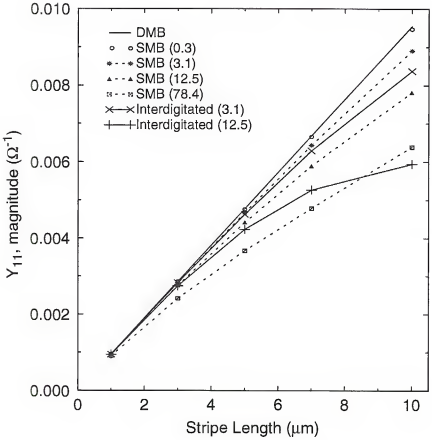


Figure 6.8 (a) The input current driving capability ($|Y_{11}|$) of different layout structures as a function of stripe length and R_{sh} indicated in parentheses.

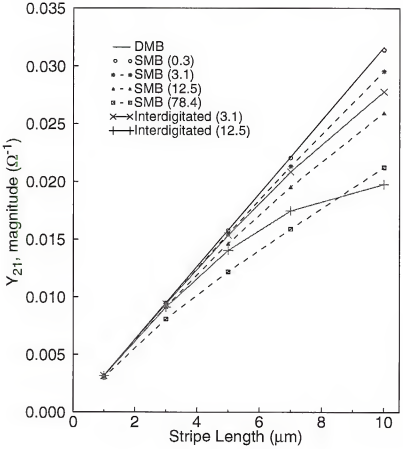


Figure 6.8 (b) The output current driving capability ($|Y_{21}|$) of different layout structures as a function of stripe length and R_{sh} indicated in parentheses

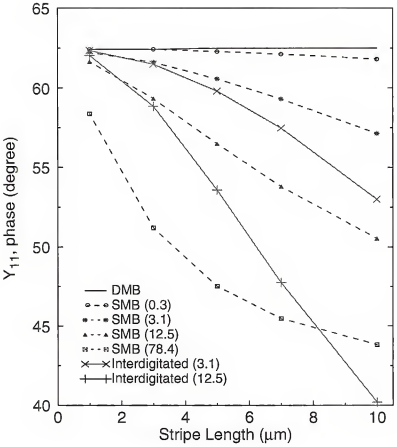


Figure 6.9 (a) The input phase delay ($\angle Y_{11}$) of different layout structures as a function of stripe length and R_{sh}.

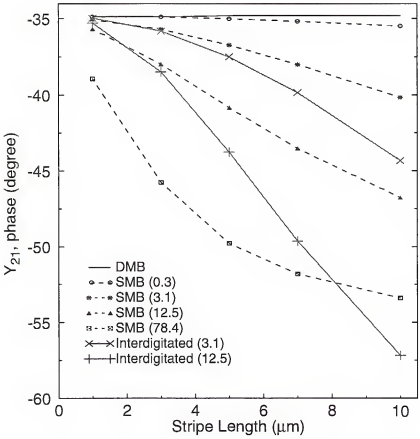


Figure 6.9 (b) The output phase delay ($\angle Y_{21}$) of different layout structures as a function of stripe length and R_{sh}.

the SMB is limited while that of the interdigitated structure keeps increasing following the increase of stripe length.

6.4. The Importance of Distributed Effects on Circuit Design

The use of different layout structures does not affect the small-signal current gain and cut-off frequency. However, there is a significant difference of Y_{11} and a minor difference in other y-parameters between various layout structures. The discrepancy requires accordant design of matching networks in an amplifier to achieve stability and power gain criterion.

Figure 6.10 illustrates the definition of reflection coefficients of a two-port network. In terms of reflection coefficients, the conditions for unconditional stability of the network at a given frequency are [Gon84]

$$|\Gamma_S| < 1\,, \tag{6-1}$$

$$|\Gamma_L| < 1\,, \tag{6-2}$$

$$|\Gamma_{IN}| = \left|S_{11} + \frac{S_{12}S_{21}\Gamma_L}{1 - S_{22}\Gamma_L}\right| < 1\,, \tag{6-3}$$

and

$$|\Gamma_{OUT}| = \left|S_{22} + \frac{S_{12}S_{21}\Gamma_S}{1 - S_{11}\Gamma_S}\right| < 1\,. \tag{6-4}$$

There are many sets of Γ_S and Γ_L that can produce stable operation. Figure 6.11 plots the stability circle of Γ_L in the Smith chart for different layout structures. The area inside the circles indicates the value of Γ_L producing unstable operation. It is found that this device is unconditionally stable for the SMB and interdigitated structures with any passive load and tends to be unconditionally stable for the DMB device unless real Γ_L approaches one. Any passive input

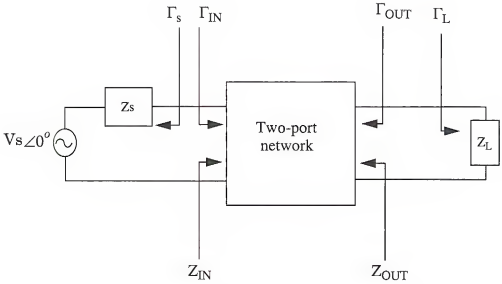
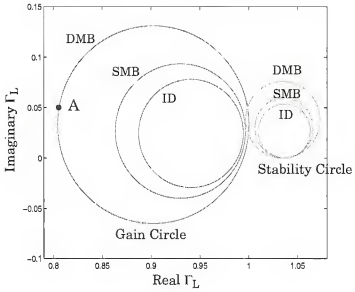
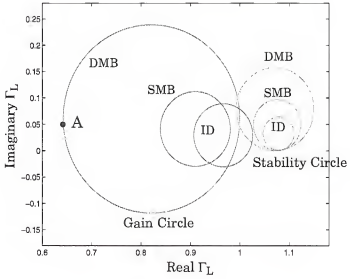


Figure 6.10 The definition of reflection coefficients of a two-port network. (After [Gon84].)



(a)



(b)

Figure 6.11 Γ_L stability circles and constant power gain circles in the Smith chart with three layout structures. The specified power gain G_p is 40. (a) 5μm stripe. (b) 10μm stripe.

impedance in the network also produces an unconditional stability for the SMB and interdigitated structures and tends to be unconditionally stable for the DMB device unless $|\Gamma_S|$ approaches one. Although all the structures tend to be unconditionally stable in this case, it is seen that the use of various layouts corresponds to different stability circles. Thus, the available Γ_S and Γ_L can be different for different layout structures in some other devices.

In microwave circuit design, power gain is one of the most important figures of merit, and the operating power gain *Gp* is independent of the source impedance. Therefore, a design procedure based on *Gp* for both unconditionally stable and potentially unstable transistors is simple and commonly used for practical design [Gon84]. *Gp* is defined as the ratio of power delivered to the load vs. power input to the network and is given by

$$Gp = \left(\frac{(1-|\Gamma_L|^2)}{\left(1-\left|\frac{S_{11}-\Delta\Gamma_L}{1-S_{22}\Gamma_L}\right|^2\right)\cdot|1-S_{22}\Gamma_L|^2} \right) \cdot |S_{21}|^2 \tag{6-5}$$

where

$$\Delta = S_{11}S_{22} - S_{12}S_{21}. \tag{6-6}$$

Both input and output characteristics of the transistor affect *Gp*. Therefore, Gp also depends on the layout structures used. A power gain circle in a Γ_L plane of the Smith chart gives the available region of Γ_L for which a specific gain can be achieved at a given frequency. The power gain circles with *Gp*=40 at a frequency of 10GHz for different layout structures are plotted in Fig. 6.11 together with the output stability circles. For both the 5 and 10 μm devices, the DMB structure allows the widest range of Γ_L that can achieve the specified *Gp*. There are more constraints for available Γ_L of the SMB device compared with

that of the DMB, and the interdigitated structure has the smallest range of available Γ_L.

As mentioned in Chapter 3 and 4, 2D device simulations can only describe devices similar to the DMB structure with uniform distributed voltage and current in the z-direction. However, 2D simulation is still one of the main tools in device design independent of the layout structures. This means the 3D distributed effects in the device and circuit design are neglected. So, how much difference will it make if these effects are not taken into account? Assuming an output matching network is designed according to 2D simulations (DMB structure) and the goal is to achieve a power gain of 40, any Γ_L in the gain circle of DMB can be used except for real $\Gamma_L \sim 1$ because of stability considerations. If a Γ_L is chosen from point A indicated in Fig. 6.10, the output matching network is simply a resistor of 433Ω and an inductor of 1.98nH in series for a 5μm stripe BJT. If this matching network is applied to a SMB device, a 1.21dB power gain loss results and a 1.75dB loss occurs in an interdigitated structure. The degradation is even worse for longer stripe transistors. Applying the matching network designed for the DMB device to a SMB device causes a 2.39dB power gain loss and a loss of 4.44dB in the interdigitated transistor. The power gain for the BJT with different layouts are summarized in Table 6.1 and 6.2.

6.5. Summary

The accuracy of the AC distributed model derived in Chapter 4 and the quasi-3D simulations have been verified by 3D device simulations in this chapter. The device characteristics dependencies on the layout structure and sheet resistance due to the presence of 3D distributed effects are also studies. It is shown that the use of different types of layout structures and geometries result in a significant difference in the input characteristics of the same

transistor unless a very low value of sheet resistance is used in the base polysilicon. A good silicided polysilicon contact greatly eliminates the debiasing and improves the small-signal current driving capability for long stripe BJTs.

The varied characteristics due to the use of different layout structures require individual design of the matching network. It is proved that applying an output matching network designed for a DMB structure to a SMB and interdigitated structure results in a significant degradation in the power gain. The stability properties can also be degraded in other cases. Therefore, the distributed effects must be taken into account in the design of SMB and interdigitated devices.

2D device simulations fail to account for the voltage and current variation in the base polysilicon region resulting in a significant error in determining AC characteristics for SMB and interdigitated structures, while 3D device simulations demand very extensive computer resources. An AC FLOODS simulation with 2100 nodes requires 1.5G of memory space running on a Sun Ultrasparc 2 computer for more than 30 hours. In contrast, the quasi-3D simulations predict the AC characteristics accurately and demand much less computer memory and simulation time making it is useful for practical device and circuit design.

Table 6.1. The power gain of a 5μm transistor with different layouts

Layout	Power Gain
DMB	40.0 (16.02dB)
SMB	30.0 (14.77dB)
Interdigitated	26.7 (14.27dB)

Table 6.2. The power gain of a 10μm transistor with different layouts

Layout	Power Gain
DMB	40.0 (16.02dB)
SMB	23.0 (13.62dB)
Interdigitated	14.4 (11.58dB)

CHAPTER 7
THE DISTRIBUTED EFFECTS ON THE AC AND TRANSIENT PERFORMANCE OF A 0.18μm MOSFET

7.1 Introduction

MOSFETs are the dominant devices in current digit circuit applications, and the applications are also being extended to radio frequency applications as the feature sizes are reduced. The conventional layout and cross section of a MOSFET with a polysilicon gate is illustrated in Fig. 7.1. The metal contacts of a gate are located at the ends of the polysilicon. The input signal needs to propagate from metal contacts through the gate to control the charges in the channel. At high frequency operation, an input signal RC delay takes place along the gate due to the distributed gate resistance and capacitance. The scaled geometries in the polysilicon gate length (channel length), thickness, and the gate oxide thickness increase the resistance and capacitance resulting in a larger RC delay. This delay becomes important if the transistor width is increased to obtain high current drive and large transconductance. This effect is especially noticeable in line drivers and output buffers used in digital circuits and low noise, high gain amplifiers employed in analog applications [Raz94] in which a wide gate is usually required.

The delay time is one of the figures of merit in measuring the performance of a MOSFET. A low gate resistance using silicided technologies is usually emphasized in device design to avoid the RC delay due to the distributed gate resistance [Mii94, Raz94]. However, the relation between the RC

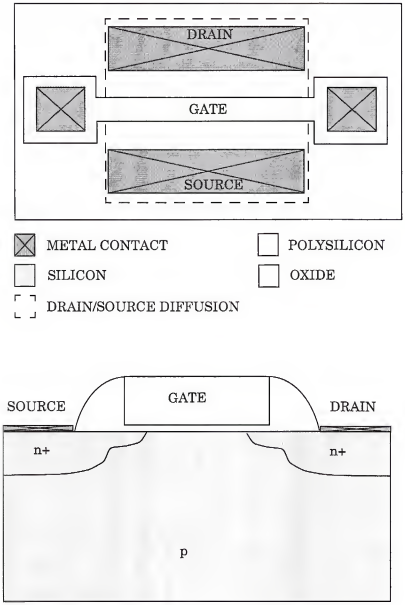


Figure 7.1. The layout and cross section of a conventional nMOSFET.

delay, device performance, and the gate sheet resistance at different frequencies is not clarified.

In this chapter, the implementation of the transient simulation scheme in FLOODS is briefly discussed. The RC delay in the transient and AC operations of a 0.18μm MOSFET with various gate sheet resistances is examined based on 3D FLOODS simulations.

7.2 Transient Simulations

The general form of the semiconductor equations is given by

$$\frac{d}{dt}q(z(t)) + f(t, z(t)) = 0 \tag{7-1}$$

where *z(t)* is the solution of the equations in nodes at time *t*, *q(z)* is the transient and *f(t,z)* is the steady-state term in the equations. The *z(t)s* and *q(z)*s for each semiconductor equation used in FLOODS are summarized in Table 7.1 where *Ec* and *Ev* are the conduction and valance bands, respectively, k_B is the Boltzmann constant, C_L is the lattice specific heat, and γ is the Fermi degenerate coefficient (refer to p. 24, [Lia94c]). The TR-BDF2 method proposed by Bank [Ban85] is used in FLOODS to discretize Eq. (7-1) in time domain.

Table 7.1

Equation	z(t)	q(z)
Poisson	potential ψ	0
Electron Continuity	electron n	n
Hole Continuity	hole p	p
Electron Energy Balance	electron temperature T_n	n(Ec+1.5$\gamma k_B Tn$)
Hole Energy Balance	hole temperature T_p	p(-Ev+1.5$\gamma k_B Tp$)
Lattice Self-Heating	lattice temperature T_L	$C_L T_L$

The total current J_{tot} in the terminals of semiconductor devices is given by

$$J_{tot} = J_n + J_p + J_{disp} \tag{7-2}$$

where J_n and J_p are the electron and hole conduction currents, respectively, and $J_{disp} = \varepsilon \frac{dE}{dt}$ is the displacement current due to the presence of space charges at the terminals, where E is the electric field and ε is the permittivity in the semiconductor material. The space charges in the terminals disappear fast due to the very short dielectric relaxation time. Therefore, the displacement current is usually not important in transient operations of contemporary semiconductor devices. Figure 7.2 shows the transient response of a diode obtained from FLOODS simulations. It is seen that the displacement current J_{disp} is only important for the first 30 femto seconds, then the carrier conduction currents J_n and J_p dominate the total current for the rest of operations. Although J_{disp} is not important in the time domain of interest, it is found that the carriers in the semiconductor are forced to response abnormally fast if J_{disp} is absent from the total current in Eq. (7-2), which results in numerical instability and significant error in the solutions of the device simulations.

7.3 Distributed Effects on the Transient Performance of MOSFETs

The simulated structure of a 0.18μm nMOSFET with a gate oxide thickness of $40\mathring{A}$ is illustrated in Fig. 7.3. The width of this transistor is 10μm and there are 2316 nodes in the simulation structure. To properly represent the location of metal contact in a real device structure, the gate metal contact in this simulation structure is attached to the face A. The input signal needs to propagate from the metal contact through the polysilicon gate to control the output characteristics in this Edge Contact (EC) structure and a RC delay

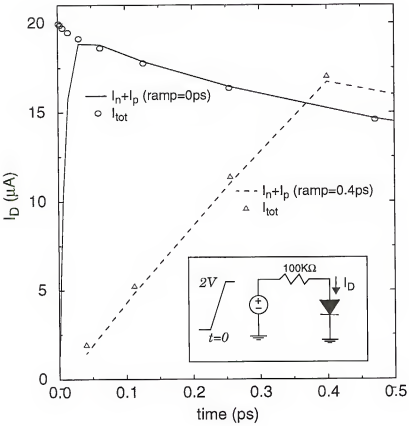


Figure 7.2. The transient current of a diode when a ramp voltage is applied at t=0. The inserted picture illustrates the circuit diagram of the simulations.

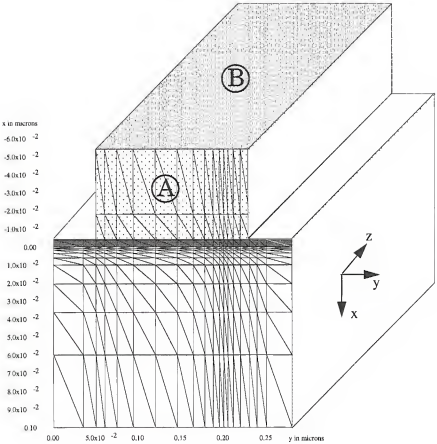


Figure 7.3. The simulation structure of a 0.18μm MOSFET. The material of the bottom cube is silicon and the cube on the top is polysilicon gate. There is a 40 Å gate oxide between silicon and gate polysilicon.

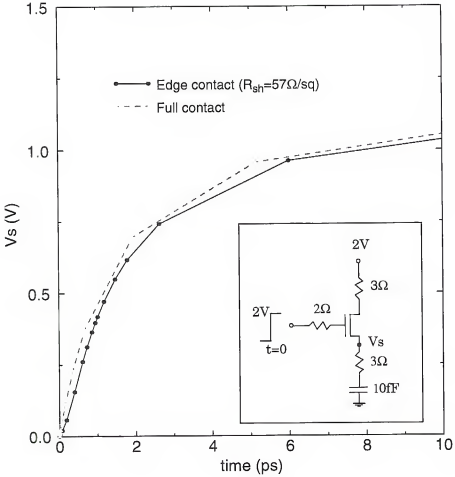


Figure 7.4. The transient source voltage in the turning on of a DRAM obtained from 3D FLOODS simulations. The inserted picture illustrates the circuit diagram of the simulations.

takes place along the signal path. The sheet resistance of the polysilicon gate is controlled by using different mobility values in the gate material which is independent of the mobilities in other materials. For comparison reasons, an artificial device in which the gate metal contact is attached to the Face B is also used in the study. Since the metal contact covers the entire polysilicon gate top surface, there is no RC delay along the gate in this Full Contact (FC) structure.

The dynamic random access memory (DRAM) cell contains a capacitor to store the charge and a MOSFET switch to read and write the stored charge. The performance of a DRAM cell depends on the switching speed of the MOSFET and the time required for the MOSFET to charge/discharge the capacitance. A delay in the input signal slows down the turning on and off of the transistor and increases the charge/discharge time. The transient characteristics of the turning on of a DRAM cell with the 0.18μm MOSFET are plotted in Fig. 7.4. In the simulations, the gate sheet resistance of the FC structure is 2Ω/sq and that of the EC structure is 57Ω/sq. A step voltage of 2V is applied to the gate metal contact at t=0. The response of the source terminal voltage is seen to significantly lag behind to the input voltage for both the FC and EC structures. However, the difference of the voltage response between FC (no gate delay) and EC (with gate delay) devices is small implying the RC delay in the gate is insignificant compared with the DRAM capacitance charging time.

The gate sheet resistance is usually lower than 57Ω/sq for contemporary MOSFET technologies, and the transistor width is normally not this wide for most of digital applications. Therefore, the distributed effects can usually be neglected in the MOSFETs for digital applications.

7.4 Distributed Effects on the AC Performance of MOSFETs

The applications of MOSFETs are being extending to the radio frequency applications as the feature sizes are reduced. A cut-off frequency of 93GHz has been reported for a 0.08μm channel MOSFET [Mii94]. The gate small-signal current increases as the operating frequency increases due to the larger displacement current passing through the gate oxide. The transistor width in the high frequency applications is usually large to achieve high current drive and large transconductance. Therefore, the AC debiasing and phase delay due to the distributed gate resistance and capacitance become important at high frequency applications.

$|Y_{11}|$ of a MOSFET increases following the increasing frequency due to the larger small-signal current flowing through the gate at higher frequencies. The phase delay of Y_{11} is also larger because of the more significant RC delay along the gate current path. The use of a larger gate sheet resistance increases the gate impedance resulting in a smaller $|Y_{11}|$ and a more significant phase delay in Y_{11}. The simulated Y_{11} of the MOSFET is shown in Fig. 7.5 as a function of frequency and gate sheet resistance at V_{GS}=3V and V_{DS}=2V. There is little phase delay for FC structure at high frequencies since there is no RC delay in the gate. Compared to the FC structure, $|Y_{11}|$ of EC structure is smaller since the effective gate impedance increases due to the longer current path. The phase delay of Y_{11} in EC device is also lagged. The difference of Y_{11} between the FC and EC structures with a R_{sh}=57Ω/sq is significant at high frequencies. If R_{sh} is reduced, the difference becomes smaller and in fact negligible if R_{sh} is less than 10Ω/sq.

As the operating frequency increases, the increasing small-signal current causes a larger debiasing along the gate which reduces the voltage across the gate and decreases the drain current. The larger input signal delay

in the gate also lags the response of the output drain current at higher frequencies. Thus, $|Y_{21}|$ decreases and phase delay of Y_{21} increases following the increase of operating frequency. Figure 7.6 shows the transconductance Y_{21} at various frequencies and R_{sh}s. $|Y_{21}|$ of FC device is independent of the frequency since there is no debiasing along the gate. A reduction of $|Y_{21}|$ is observed at high frequencies for the EC structures. Larger R_{sh} causes more significant decrease due to a bigger small-signal ohmic voltage drop along the gate. The phase delay of Y_{21} for the EC devices is also larger at high frequencies compared with that of the FC device due to the additional gate RC delay. However, the difference between the FC and EC structures is insignificant if the gate sheet resistance is smaller than 10Ω/sq. The analysis on the Y_{12} and Y_{22} also reveals the distributed effect degrades the characteristics of devices at high frequencies for large gate sheet resistance and becomes negligible if R_{sh} is less than 10Ω/sq.

Therefore, a MOSFET with a large gate sheet resistance suffers from the distributed effect at high frequencies and the performance degrades. The silicided gate technologies eliminate the distributed effect and improve the performance. If the gate sheet resistance is less than 10Ω/sq, this effect can be neglected in the device and circuit design.

7.4 Summary

The distributed effects on the performance of an advanced 0.18μm nMOSFET have been studied based on 3D device simulations. It has been shown that the RC delay due to the distributed gate resistance and capacitance contributes little to the delay time of a DRAM cell even if a large gate sheet resistance is used. Therefore, the performance of MOSFETs in digital applications is fairly insensitive to silicided gate technologies. The distributed effects

are significant for MOSFETs with large gate sheet resistance in radio frequency applications. The use of silicided gate technologies is necessary to improve the performance of high speed MOSFETs. However, if a reasonably low polysilicon gate sheet resistance is utilized, the distributed effects and RC delay in the gate can be neglected in the radio frequency device and circuit design.

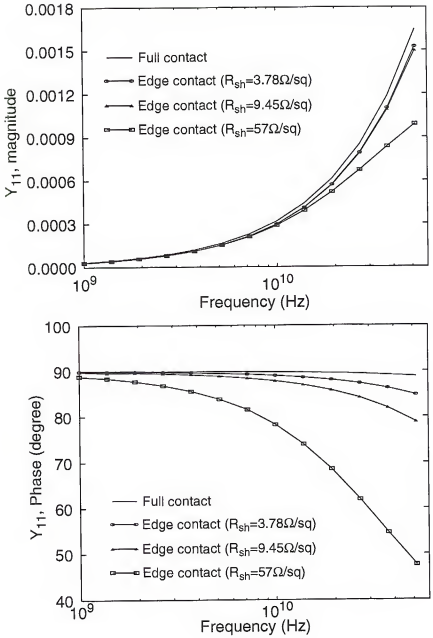


Figure 7.5. The simulated Y_{11} as a function of frequency and R_{sh}.

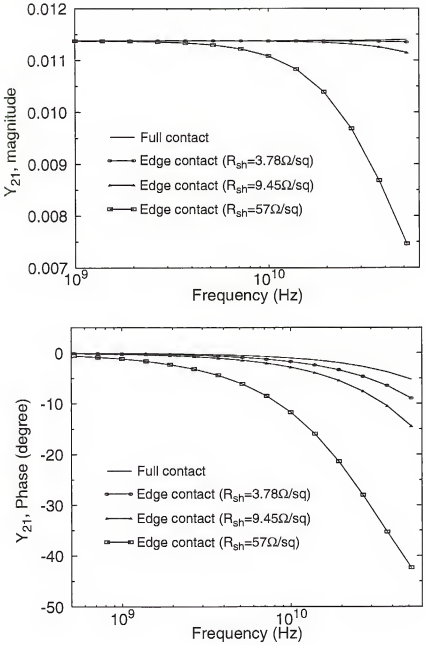


Figure 7.6. The simulated Y_{21} as a function of frequency and R_{sh}.

CHAPTER 8
SUMMARY AND FUTURE WORK

8.1 Summary

The voltage and current in the base region of a BJT and in the polysilicon gate region of a MOSFET exhibit the distributed effects. This distributed effect is significant when a device operates at high levels of current injection or at high speed. Doing device and circuit design without taking into account this effect can result in degraded device and circuit performance.

Distributed effects in the polysilicon and intrinsic base regions of stripe BJTs with double polysilicon technology were studied in detail in this dissertation. Analytical models describing the DC and AC voltage and current distributions in the devices were presented in Chapters 3 and 4, respectively. This quasi-3D model properly predicted the input characteristics of a single metal base contact BJT with double polysilicon technologies. A quasi-3D device simulation method was also demonstrated in chapter 4 to accurately describe the 3D effects using 2D device simulations, which enables engineers to design devices and circuits efficiently.

The influences of the distributed effect on various types of BJT layout structures, geometry sizes, and polysilicon contact resistances were studied in Chapters 5 and 6 for DC and AC operations, respectively. It was shown that very low resistance polysilicon is necessary for high performance BJTs. Therefore, a silicided polysilicon process is one of the key factors in determining the device performance of a BJT. The use of different layout structures and geome-

try sizes also results in significant differences in input and output characteristics of BJTs. The linear relation between current drive and geometry sizes is invalid in devices with long stripe operating at high biases or high frequencies. A design without taking into account the distributed effect was shown to result in a poor circuit performance.

The distributed effect was also observed in an advanced MOSFET. This effect significantly degrades the transconductance and delay the output response if a large gate sheet resistance is used for high frequency applications. The silicided polysilicon technologies are necessary to improve the performance of high speed MOSFETs. However, the effect is less important on the high frequency performance of MOSFETs if an adequate polysilicon gate sheet resistance is utilized. The results also showed the distributed effect does not significantly affect the MOSFET performance in digital applications even if a large gate sheet resistance is utilized.

8.2 Future Work

The following subjects are suggested to continue the applications of distributed effects:

(1) Implementing the automatic quasi-3D simulation capability in FLOODS. Once the 3D device information are given, all the device parameters and coefficients in the model can be determined from 2D device simulations. The effective lengths are then computed from the analytical equations and feed to the simulator for quasi-3D simulations. This procedure can be included in a macro function of FLOODS to automatically perform the quasi-3D simulations.

(2) Including the automatic extraction function for base resistance in FLOODS. Base resistance is one of the most important and most difficult parameters in BJTs to extract. All the device parameters and coefficients required in the

model to compute the base resistance can be obtained from 2D device simulations. Therefore, a post-processor reads the 2D simulation results can compute the base resistance directly.

(3) Including the distributed effects in a compact model. A circuit simulator is the main tool in current integrated circuit design. The accuracy of the simulation results depends on the accuracy of the compact device models used in the simulator. AC base resistance is one of the most important parameters in the compact model. The impedance semicircle method based on the simple lumped circuit model [Fus95] is one way to extract the AC base resistance. Figure 8.1 shows an impedance semicircle in a complex plane obtained from 3D simulations with a 5μm stripe length SMB BJT similar to the structure used in chapter 5. The base resistance is extracted from the intersection of the semicircle and the real axis at high frequencies. However, the input impedance obviously cannot be expressed by a circle equation due to the input distributed resistance and capacitance. Therefore, the inclusion of this effect in the compact models is important for high speed circuit designs.

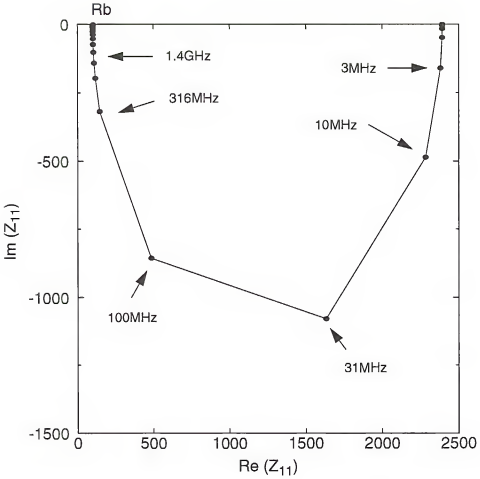


Figure 8.1. Locus of input impedance in complex plane.

REFERENCES

Ago94 V. M. Agostinelli T. J. Bordelon, X. Wang, K. Hasnat, C. F. Yeap, D. B. Lemersal, A. F. Tasch, and C. M. Maziar, "Two-dimensional energy-dependent models for the simulation of substrate current in submicron MOSFET's," IEEE Trans. Electron Devices, vol. ED-41, no. 10, pp. 1784-1795, 1994.

Auf96 K. Aufinger, J. Bock, T. F. Meister, and J. Popp, "Noise characteristics of transistors fabricated in an advanced silicon bipolar technology," IEEE Trans. Electron Devices, vol. ED-43, pp. 1533-1538, 1996.

Ban85 R. E. Bank, W. M. Coughran, W. Fichtner, D. J. Rose, and R. K. Smith, "Transient simulation of silicon devices and circuits," IEEE Trans. Electron Devices, vol. ED-32, pp. 1992-2007, 1985.

Bur89 J. N. Burghartz, S. R. Mader, B. S. Meyerson, B. J. Ginsberg, J. M. Stork, C. Stanis, and J. Y. C. Sun, "Self-aligned bipolar npn transistor with 60nm epitaxial base," IEDM Tech. Dig., pp. 229-232, 1989.

Che93 D. Chen, Z. Yu, K. C. Wu, R. Goossens, and R. Dutton, "Dual energy transport model with coupled lattice and carrier temperature," Simulation of Semiconductor Devices and Processes, vol. 5, Sep. 1993.

Chi91 T. Y. Chiu, G. M. Chin, M. Y. Lau, R. C. Hanson, M. D. Morris, K. F. Lee, and T. Y. Liu, "The design and characterization of nonoverlapping super self-aligned BiCMOS technology," IEEE Trans. Electron Devices, vol. ED-38, pp. 141-149, 1991.

Cho90 Hanjin Cho and D. E. Burk, "A simple model for distributed base impedance with AC verification using S-parameters measurements," IEEE Bipolar Circuits and Technology Meeting, Minneapolis, Minn., pp. 106-109, 1990.

Cho93 E. F. Chor and L. S. Tan, "Analytical and numerical studies of the dependence of high frequency performance on the use of a polysilicon emitter in bipolar transistors," Solid-State Electronics, vol. 36, no. 4, pp. 553-562, 1993.

Com90 J. H. Comfort, G. L. Patton, J. D. Cressler, W. Lee, and E. F. Crabbe, "Profile leverage in a self-aligned epitaxial Si or SiGe base bipolar technology," IEDM Tech. Dig., pp. 21-24, 1990.

Coo83 R. K. Cook, "Numerical simulation of hot-carrier transport in silicon bipolar transistors," IEEE Trans. Electron Devices, vol. 30, no. 9, pp. 1103-1110, 1983.

Coo71 H. F. Cooke, "Microwave transistors: theory and design," Proceeding of the IEEE, vol. 59, no. 8, pp. 1163-1181, 1971.

Dub94 E. Dubois, P. H. Bricout, and E. Robilliart, "Accuracy of series resistances extraction schemes for polysilicon bipolar transistors," IEEE Bipolar Circuits and Technology Meeting, Minneapolis, Minn., pp. 148-151, 1994.

Elt82 A. A. Eltoukhy and D. J. Roulston, "The role of the interfacial layer in polysilicon emitter bipolar transistor," IEEE Trans. Electron Devices, vol. ED-29, pp. 1862-1869, 1982.

Fos80 J. G. Fossum and M. A. Shibib, "A minority carrier transport model for polysilicon contacts to silicon bipolar devices, including solar cells," IEDM Tech. Dig., pp. 280-283, 1980.

Fuj91 H. Fujimaki, K. Suzuki, Y. Emumera, and K. Akahane, "A novel self-aligned epitaxial base transistor," IEEE Bipolar Circuits and Technology Meeting, Minneapolis, Minn., pp. 59-62, 1991.

Fus92 T. Fuse, "A physically based base pushout model for submicrometer BJT's in the presence of velocity overshoot," IEEE Trans. Electron Devices, vol. ED-39, pp. 396-403, 1992.

Fus95 T. Fuse and Y. Sasaki, "An analysis of small-signal and large signal base resistance for submicrometer BJT's," IEEE Trans. Electron Devices, vol. ED-42, no. 3, pp. 534-539, 1995.

Fus95b K. Fushinobu, A. Majumdar, K. Jijikata, "Heat generation and transport in submicron semiconductor devices," J. Heat Transfer, vol. 117, pp. 25-31, 1995.

Gau79 Santosh P. Gaur, "Performance limitations of silicon bipolar transistors," IEEE Journal of Solid-State Circuits, vol. SC-14, no. 2, pp. 337-343, Apr. 1979.

Gho65 H. N. Ghosh, "A distributed model of the junction transistor and its application in the prediction of the emitter-base diode characteristics, base impedance, and pulse response of the device," IEEE Trans. Electron Devices, vol. 12, pp. 513-531, 1965.

Gon84 G. Gonzalez, "Microwave transistor amplifiers, analysis and design," Prentice-Hall, Englewood Cliffs, NJ, 1984.

Gra76 J. Graul, A. Glasl, and H. Murrnmann, "High-performance transistors with arsenic-implanted polysilicon emitters," IEEE J. Solid-State Circuits, vol. 11, pp. 491-494, 1976.

Gra79 H. C. de'Graaff and J. G. de Groot, "The SIS tunnel emitter-A theory for emitters with thin interface layers," IEEE Trans. Electron Devices, vol. ED-26, pp. 1771-1776, 1979.

Hau64 J. R. Hauser, "The effects of distributed base potential on emitter-current injection density and effective base resistance for stripe transistor geometries," IEEE Trans. Electron Devices, vol. ED-11, pp. 238-242, 1964.

Ich89 H. Ichino, N. Ishihara, M. Suzuki, and S. Konaka, "18GHz 1/8 dynamic frequency divider using Si bipolar technologies," IEEE Journal of Solid-State Circuits, vol. SC-24, pp. 1723, 1989.

Ino94 K. Inou, S. Matsuda, H. Nakajima, N. Sugiyama, S. Imai, Y. Kawaguchi, K. Yamada, and H. Iwai, "52GHz epitaxial base bipolar transistor with high Early voltage of 26.5V with box-like base and retrograded collector impurity profiles," IEEE Bipolar Circuits and Technology Meeting, pp. 217-220, 1994.

Ker92 M. Kerber, E. Bertagnolli, R. Mahnkoph, J. Popp, A. Felder, H. M. Rein, A. Weisgerber, and H. Klose, "A high performance BiCMOS process featuring 40GHz/21ps," IEDM Tech. Dig., pp. 449-452, 1992.

Kir62 C. T. Kirk, "A theory of transistor cutoff frequency (f_T) fall-off at high current density," IEEE Trans. Electron Devices, vol. ED-9, pp. 164, 1962.

Kiz93 I. C. Kizilyalli, T. E. Ham, K. Singhal, J. W. Kearney, W. Lin, and M. J. Thoma, "Predictive worst case statistical modeling of 0.8-μm BICMOS bipolar transistors: A methodology based on process and mixed device/circuit level simulators," IEEE Trans. Electron Devices, vol. ED-40, pp. 966-972, 1993.

Kum93 M. J. Kumar, A. D. Sadovnikov, and D. J. Roulston, "Collector design tradeoffs for low voltage applications of advanced bipolar transistors," IEEE Trans. Electron Devices, vol. ED-40, no. 8, pp. 1478-1483, 1993.

Kum94 M. J. Kumar and D. Roulston, "Optimum collector epi-thickness of advanced bipolar transistors for high speed and high current operation," Solid-State Electronics, vol. 37, no. 11, pp. 1885-1887, 1994.

Lau85 S. Laux, "Techniques for small-signal analysis of semiconductor devices," IEEE Trans. on CAD, vol. 4, no. 4, pp. 472-481, 1985.

Lia94a M. Liang, "Composite thermodynamic system for advanced semiconductor device simulation," Ph. D. dissertation, University of Florida, Gainesville, 1994.

Lia94b M. Liang and M. E. Law, "Influence of Lattice Self-Heating and Hot-Carrier Transport on Device Performance," IEEE Trans. Electron Devices, vol. ED-41, no. 12, pp. 2391-2398, 1994.

Lia94c M. Liang and M. E. Law, "An Object-Oriented Approach to Device Simulation-FLOODS," IEEE Trans. on CAD, vol. 13, no. 10, pp. 1235-1240, 1994.

May92 K. Mayaram and D. O. Pederson, "Coupling algorithms for mixed-level circuit and device simulation," IEEE Trans. on CAD, vol. 11, no. 8 pp. 1003-1012, 1992.

Mii94 Y. Mii, S. Rishton, Y. Taur, D. Kern, T. Lii, K. A. Jenkins, D. Quinlan, and M. Polcari, "Experimental high performance sub-0.1μm channel nMOSFET's," IEEE Electron Device Letter, vol. 15, no. 1, pp. 28-30, 1994.

Nak80 H. Nakashiba, "An advanced PSA technology for high-speed bipolar LSI," IEEE Trans. Electron Devices, vol. ED_27, pp. 1390-1394, 1980.

Nak95 T. Nakamura and H. Nishizawa, "Recent progress in bipolar transistor technology," IEEE Trans. Electron Devices, vol. 42, no. 3, pp. 390-398, 1995.

Nin80 T. H. Ning and R. D. Issac, "Effect of emitter contact on current gain of silicon bipolar devices," IEEE Trans. Electron Devices, vol. ED-27, pp. 2051-2055, 1980.

Oka88 H. Oka, S. Satoh, and N. Nakayama, "Bipolar circuit simulation system using a two-dimensional device simulator," Simulation of Semiconductor Devices and Processes, vol. 3, pp. 127-138, 1988.

Par86 H. K. Park, K. Boyer, A. Tang, C. Clawson, S. Yu. T. Yamaguchi, and J. Sachitano, "High speed polysilicon emitter-base bipolar transistor," IEEE Electron Device Lett., vol. EDL-7, pp. 658, 1986.

Pos92 I. R. C. Post, P. Ashburn, and G. Wolstenholme, "Polysilicon emitters for bipolar transistors: A review and re-evaluation of theory and experiment," IEEE Trans. Electron Devices, vol. 39, pp. 1717-1731, 1992.

Pri61 R. L. Pritchard, "Transistor internal parameters for small-signal representation," Proc. IRE, vol. 49, pp. 725, 1961.

Raz94 B. Razavi, R. H. Yan, and K. F. Lee, "Impact of distributed gate resistance on the performance of MOS devices," IEEE Trans. Circuits and Systems, vol. ED-41, pp. 750-754, 1994.

Rhe95 H. S. Rhee, S. Lee, and B. R. Kim, "D.C. and A.C. current crowding effects model analysis in bipolar junction transistors using a new extraction method," Solid-State Electronics, vol. 38, no. 1, pp. 31-35, 1995.

Rol88 J. G. Rollins and J. Choma, "Mixed-mode PISCES-SPICE coupled circuit and device solver," IEEE Trans. on CAD, vol. 7, no. 8, pp. 862-867, 1988.

Sat90 H. Satake and T. Hamasaki, "Low-temperature (77K) BJT model with temperature dependences on the injected condition and base resistance," IEEE Trans. Electron Devices, vol. 37, pp. 1688-1697, 1990.

Sch91 M. Schroter, "Simulation and modeling of the low-frequency base resistance of bipolar transistors and its dependence on current and geometry," IEEE Trans. Electron Devices, vol. ED-38, pp. 538-544, 1991.

Set75 J. Y. W. Seto, "The electrical properties of polycrystalline silicon films," J. Appl. Phys., vol. 46, pp. 5247-5254, 1975.

Shr90 M. Shur, "Physics of semiconductor devices," Prentice-Hall, Englewood Cliffs, NJ, 1990.

Sol79 P. M. Solomon and D. D. Tang, "Bipolar circuit scaling," ISSCC Dig. Tech. Papers, pp. 86-87, 1979.

Spi65 Philip Spiegel, "Transistor base resistance and its effect on high speed switching," Solid State Design, pp. 15-18, Dec. 1965.

Str62 R. Stratton, "Diffusion of hot and cold electrons in semiconductor barriers," Phys. Rev., vol. 126, no. 9, pp. 2002-2014, 1962.

Suz85 M. SuZuki, K. Hagimoto, H. Ichino, and S. Konaka, "A 9-GHz frequency divider using Si bipolar super aligned process technology," IEEE Electron Device Lett., vol. EDL-6, pp. 181, 1985.

Sze81 S. M. Sze, "Physics of semiconductor devices," Wiley, New York, 1981.

Taf91 R. C. Taft, J. D. Hayden, and C. D. Gunderson, "Optimization of two-dimensional collector doping profiles for submicron BiCMOS technologies," IEDM Tech. Dig., pp. 869-872, 1991.

Tan79 D. D. Tang and P. M. Solomon, "Bipolar transistor design for optimized power-delay logic circuits," IEEE Journal of Solid-State Circuits, vol. SC-14, no. 4, pp. 679-684, 1979.

Tan82 D. D. Tang, P. M. Solomon, T. K. Ning, R. D. Isaac, and R. E. Burger, "1.25μm deep-groove-isolated self-aligned bipolar circuits," IEEE Journal of Solid-State Circuits, vol. SC-17, no. 2, pp. 925-931, 1982.

Tor91 V. de la Torre, J. Foerstner, B. Lojek, K. Sakamoto, S. L. Sundaram, N. Tracht, B. Vasquez, and P. Zdebel, "MOSAIC V-A very high performance bipolar technology," IEEE Bipolar Circuits and Technology Meeting, pp. 21-24, 1991.

Ver91 M. P. J. G. Versleijen, "Distributed high frequency effects in bipolar transistors," IEEE Bipolar Circuits and Technology Meeting, pp. 85-88, 1991.

War95 J. D. Warnock, "Silicon bipolar device structures for digital applications: technology trends and future directions," IEEE Trans. Electron Devices, vol. ED-42, no. 3, pp. 377-389, 1995.

Wen92a J. Weng, "Transit time of fast bipolar transistors at high collector-current densities," Solid-State Electronics, vol. 35, no. 4, pp. 599-610, 1992.

Wen92b J. Weng, K. Ehinger, and T. F. Meister, "Collector optimization for high-speed bipolar transistors," IEEE Trans. Electron Devices, vol. ED-39, no. 5, pp. 1240-1242, 1992.

Whi70 M. H. White and M. O. Thurston, "Characterization of microwave transistors," Solid-State Electronics, vol. 13, pp. 523-542, 1970.

Wol86 S. Wolf and R. N. Tauber, "Silicon processing for the VLSI era, Volume 1-process technology," Lattice press, Sunset Beach, CA, 1986.

Yam93 T. Yamaguchi, S. Uppili, J. S. Lee, G. H. Kawamoto, T. Dosluoglu, and S. Simpkins, "Process and device characterization for a 30-GHz f_T submicrometer double poly-si bipolar technology using BF_2-implanted base with rapid thermal process," IEEE Trans. Electron Devices, vol. ED-40, pp. 1484-1495, 1993.

Yam94 K. Yamano, H. Fujimaki, H. Yokouchi, K. Ohshima, and K. Suzuki, "A high speed bipolar transistor using 2-step epitaxial base technology," IEEE Bipolar Circuits and Technology Meeting, pp. 64-67, 1994.

Yu84 Z. Yu, B. Ricco, and R. W. Dutton, "A Comprehensive analytical and numerical model of polysilicon emitter contacts in bipolar transistors," IEEE Trans. Electron Devices, vol. ED-31, pp. 773-784, 1984.

Yua88 J. S. Yuan, "A physics-based current-dependent base resistance model for advanced bipolar transistors," IEEE Trans. Electron Devices, vol. 35, pp. 1055-1062, 1988.

Zde87 P. J. Zdebel, R. J. Balda, B. V. Hwang, V. del la Torre, and A. Wagner, "MOSAIC III-A high performance bipolar technology with advanced self-aligned devices," IEEE Bipolar Circuits and Technology Meeting, pp. 172-175, 1987.

Zie58 A. van der Ziel, "Noise in junction transistors," Proceedings of the IRE, pp. 1019-1038, June1958.

BIOGRAPHICAL SKETCH

Ming-Yeh Chuang was born in Taiwan in 1964. He received the B.S. degree from Tamkang University, Taiwan, in 1986, and the M.S. degree from the University of Florida in 1993, both in electrical engineering.

Since 1993, he has been working towards the Ph.D. degree in electrical engineering at the University of Florida. His research interests are in simulating and modeling physical processes in semiconductor devices, specifically high frequency transistors. His work involves the development of the general purpose device simulator FLOODS for analysis of advanced semiconductor devices.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Mark E. Law, Chairman
Professor of Electrical and Computer Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Jerry G. Fossum
Professor of Electrical and Computer Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Sheng S. Li
Professor of Electrical and Computer Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Robert M. Fox
Associate Professor of Electrical and Computer Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Kenneth K. O
Assistant Professor of Electrical and Computer Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Loc Vu-Quoc
Associate Professor of Aerospace Engineering,
Mechanics and Engineering Science

This dissertation was submitted to the Graduate Faculty of the College of Engineering and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August 1997

Winfred M. Phillips
Dean, College of Engineering

Karen A. Holbrook
Dean, Graduate School